

Métodos Básicos de Previsão

O que é previsão?

"Prediction is very difficult, especially if it's about the future."
Niels Bohr, laureado com Nobel em Física

Segundo o Dicionário do Aurélio (FERREIRA, 1975) a palavra previsão é um substantivo feminino que significa ato ou efeito de prever; antevisão; estudo ou exame feito com antecedência. Dessa forma podemos conceituar **previsão** como a estimação do valor de uma variável (ou conjunto de variáveis) em algum momento futuro. Aqui consideraremos alguns métodos de previsão.

Um exercício de previsão é geralmente realizado para fornecer uma ajuda à tomada de decisão e no planejamento do futuro.

Tipicamente se pudermos prever o como será o futuro podemos modificar o nosso comportamento agora para ficar numa posição melhor que aquela que por outro lado seria, quando o futuro chegar. Uma previsão adequada deve dar suporte a uma decisão minimizadora de risco por parte dos tomadores de decisão (HARRISON e STEVENS, 1976). E é essencial, portanto, para o planejamento e administração de recursos. As aplicações para a previsão incluem:

- Controle de estoque/ planejamento da produção - prever a demanda para um produto habilita-nos a controlar o estoque de matéria prima e produtos acabados, planejar a programação de produção, etc.
- Política de investimento – prever informações financeiras tais como taxas de juros, taxas de câmbio, preços de ações, o preço do ouro, etc. Esta é uma área que ninguém ainda desenvolveu uma técnica de previsão (consistentemente acurada ou no mínimo se contaram ou não a alguém!).
- Política econômica – prever informação econômica tal como o crescimento da economia, desemprego, taxa de inflação, etc. é vital para ambos o governo e negócios no planejamento do futuro.

Tipos de problemas

Um modo de classificar os problemas de previsão é considerar a escala de tempo envolvida na previsão, i.é., quão longe no futuro estamos tentando prever. Curto, médio e longo prazos são as categorias usuais mas o significado real de cada um variará de acordo com a situação que está sendo estudada, por exemplo, na previsão de demanda de energia para construir estações de força, 5-10 anos seria curto prazo e 50anos seria longo prazo, enquanto que na previsão de demanda de consumo em muitos negócios situações de até 6 meses seria curto prazo e mais do que dois anos longo prazo. A tabela 01 abaixo mostra a escala de tempo associada com decisões de negócios.

Métodos de previsão

A razão básica para a classificação da tabela 01 é que diferentes métodos de previsão se aplicam em cada situação, p. ex., um método de previsão que é apropriado para previsão de vendas no próximo mês (previsão de curto-prazo) provavelmente seria não apropriado para previsão de vendas daqui a cinco anos (previsão de longo prazo). Em particular note aqui que o uso de números (dados) para os quais as técnicas quantitativas são aplicadas tipicamente variando de previsão muito alta para a de curto prazo e muito baixa para a previsão de longo prazo quando estamos tratando com situações de negócios. Algumas áreas podem envolver previsões de curto, médio e longo prazos, como mercado de ações e previsões do tempo.

¹ Previsão é muito difícil, especialmente se ela for acerca do futuro.

Os métodos de previsões podem também ser classificados em várias categorias diferentes:

- *Métodos qualitativos* – onde não há modelo matemático formal, frequentemente porque os dados disponíveis não são imaginados como representativos do futuro (previsão de longo prazo). Há vários métodos qualitativos, como a abordagem Delphi ou opinião de especialista, pressupostos gerenciais, bem como as pesquisas de mercado, dados externos ou sondagens e pesquisa (MUN, 2010).
- *Métodos quantitativos* – onde os dados históricos das variáveis de interesse são disponíveis – estes métodos são baseados numa análise dos dados históricos relativos à série temporal de uma variável específica de interesse e possivelmente outras relacionadas à série temporal e também examina as relações de causa e efeito da variável com outras variáveis relevantes.

Tabela 01

Escala de tempo	Tipo de Decisão	Exemplos
Curto prazo até 3-6 meses	Operacional	Controle de estoque, planejamento da produção, distribuição
Prazo médio de 3-6 meses até 2 anos	Tática	Leasing de planta e equipamento, mudanças de emprego
Longo prazo acima de 2 anos	Estratégica	Pesquisa e Desenvolvimento, Aquisições e fusões, variações de produto

Na parte *quantitativa*, as abordagens de previsão geralmente caem numa dessas categorias:

- *Séries Temporais* – Realiza a análise de séries temporais sobre a forma padrão passada dos dados para, a partir daí projetar os resultados futuros. Isto funciona bem para situações estáveis em que as condições são esperadas permanecerem as mesmas. Como exemplo, podemos citar as receitas em diferentes anos, os índices de inflação, as taxas de juros, a participação no mercado, taxas de falhas, os preços nominais mensais da arroba do boi gordo divulgados pelo CEPEA, e assim por diante.
- *Regressão* – Prevê os resultados futuros usando relações passadas entre uma variável de interesse (variável dependente) e várias outras variáveis que podem influenciá-la (variáveis independentes). Isso funciona bem para situações em que se precisam identificar os diferentes efeitos das diferentes variáveis. Nesta categoria se inclui a *regressão linear múltipla*.
- *Simulação* – Aleatoriamente gera muitos cenários para um modelo de previsão dos resultados possíveis. Este método funciona bem onde não se tem dados históricos, mas se pode construir o modelo da situação para analisar o seu comportamento. Atualmente a simulação de Monte Carlo é usada largamente nestas situações e os suplementos do MS-Excel, Crystal Ball e Risk Simulator são popularmente conhecidos para automatizar esta tarefa.

Diferentes Tipos de Técnicas de Previsão

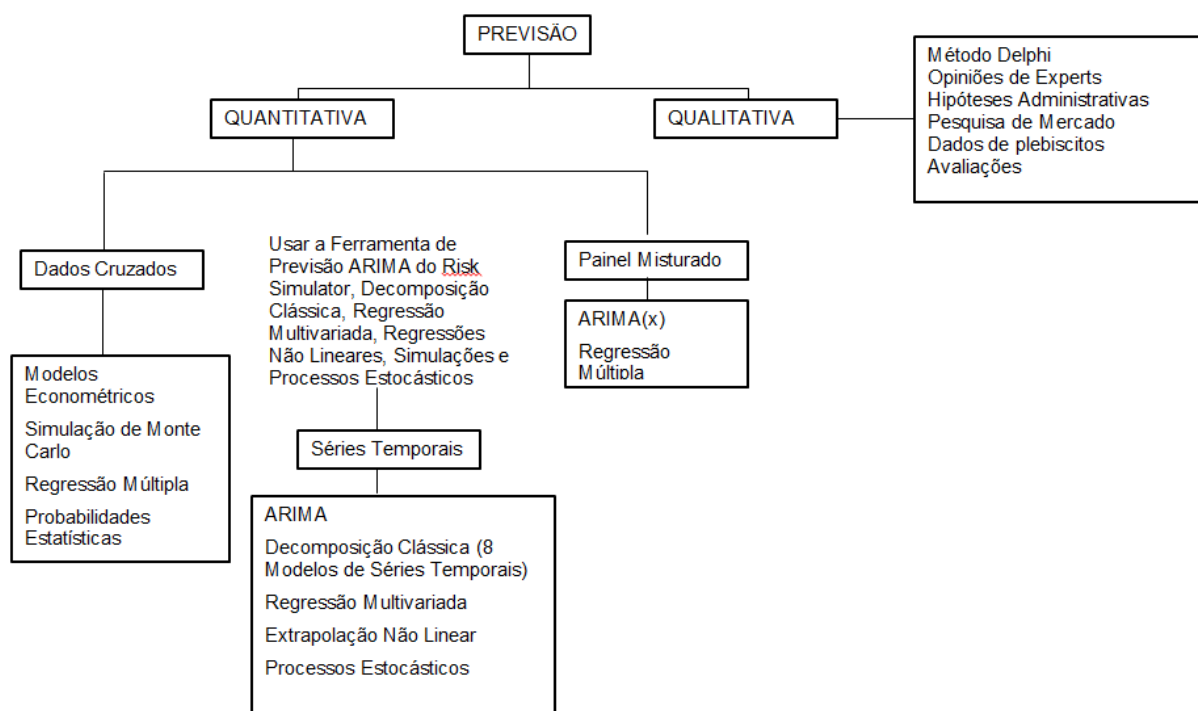


Figura 1 – Metodologias mais comuns para a Previsão.

Fonte: MUN, J. **Modeling Risk: Applying Monte Carlo Simulation, Real Options Analysis, Stochastic Forecasting and Portfolio Optimization**. 2ª ed., New York: John Wiley & Sons, 2010, página 261, adaptado.

Como mostrado na Figura 1, existem vários métodos para se realizar a previsão. Os mais comuns são:

1. ARIMA (modelo de média móvel integrada autorregressiva) de Box-Jenkins
2. AutoARIMA
3. Econometria básica
4. Autoeconometria
5. Lógica difusa combinatória
6. Distribuições personalizadas
7. GARCH (modelo autoregressivo à heteroscedasticidade condicional generalizado)
8. Curvas J
9. Cadeias de Markov
10. Máxima verossimilhança
11. Rede Neural
12. Regressão multivariada
13. Extrapolação não linear
14. Curvas S
15. *Spline* cúbico
16. Previsão estocástica

17. Análise da série temporal

18. Linhas de tendência

A Metodologia de Previsão por análise de séries temporais (*time-series forecasting*).

"O tempo diz o que a razão não pode dizer."
René Descartes

A análise de séries temporais se aplica nos casos em que há um padrão persistente ou sistemático no comportamento da variável, que é possível de captar através de uma representação paramétrica (PINDYCK e RUBENFIELD, 1991). Alguns métodos clássicos são projetados para funcionarem melhor para certos tipos de dados:

- Dados sazonais (aumentando ou diminuindo num padrão recursivo regular no decorrer do tempo)
- Dados com tendência (aumentando ou diminuindo consistentemente no decorrer do tempo)

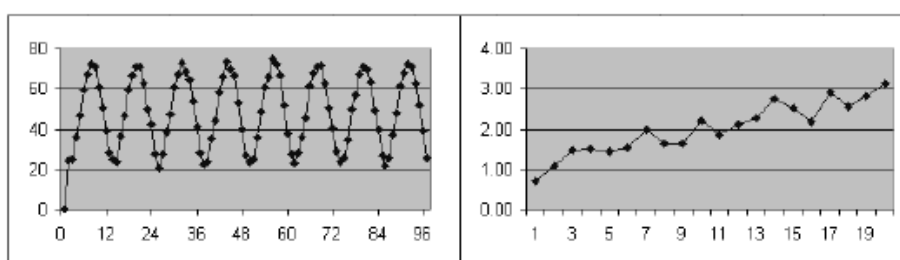


Figura 2 – Dados sazonais e dados com tendência.

Fonte: CB Predictor – User Manual – Selecionando métodos de previsão de séries temporais. P. 3-8.

Segundo (MORETTIN & TOLOI, 2004), pode-se entender a **tendência** (T) como o movimento persistente nos dados em uma dada direção e a sazonalidade (S) como o comportamento regular assumido pela série em algum subperíodo. O erro aleatório (ε) leva em consideração movimentos esporádicos e irregulares presentes na série.

A previsão de séries temporais assume que os dados históricos sejam uma combinação de um formato padrão e de algum erro aleatório. Sua meta é isolar o formato do erro entendendo o nível, a tendência e a sazonalidade do formato. Pode-se medir o erro usando uma medida estatística para descrever quão bem um formato reproduz os dados históricos e estimar quão acuradamente ele projeta os dados no futuro.

A Figura 3 lista os oito modelos clássicos de *série temporal*, separados por **sazonalidade** e **tendência**. Por exemplo, se a variável de dados **não** possuir tendência ou sazonalidade, um modelo de *média móvel simples* ou de *suavização exponencial simples* seria suficiente. No entanto, se existir sazonalidade, mas nenhuma tendência discernível estiver presente, um modelo *sazonal aditivo* ou *sazonal multiplicativo* seria melhor, e assim por diante.

	Não Sazonalidade	Com Sazonalidade
Nenhuma Tendência	<i>Média Móvel Simples</i>	<i>Sazonal Aditivo</i>
	<i>Suavização Exponencial Simples</i>	<i>Sazonal Multiplicativo</i>
Com Tendência	<i>Média Móvel Dupla</i>	<i>Aditivo de Holt-Winter</i>
	<i>Suavização Exponencial Dupla</i>	<i>Multiplicativo de Holt-Winter</i>

Figura 3 – Os oito métodos clássicos de série temporal.

Fonte: MUN, J. **Modeling Risk: Applying Monte Carlo Simulation, Real Options Analysis, Stochastic Forecasting and Portfolio Optimization**. 2ª ed., New York: John Wiley & Sons, 2010. Página 264.

O que se pretende sempre é testar cada um destes métodos clássicos e classifica-los de acordo com o erro. O método com o erro mais baixo é o melhor método.

Existem dois tipos de métodos sazonais: *aditivo* e *multiplicativo*. A sazonalidade **aditiva** tem um padrão estacionário de amplitude, e a sazonalidade **multiplicativa** tem um padrão de amplitude crescendo ou decrescendo no decorrer do tempo.

A Figura 4 mostra as diferentes curvas de sazonalidades

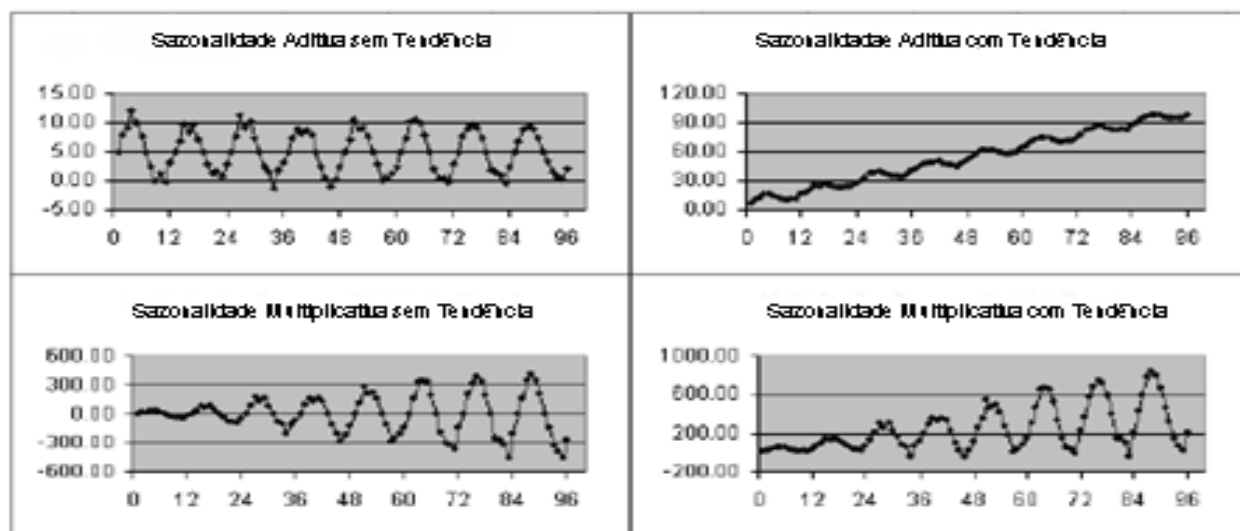


Figura 4 – Dados sazonais e dados com tendência.

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Existem duas técnicas principais de previsão de séries temporais:

1. Suavização de Não Sazonalidade que estima uma tendência ou não removendo dados extremos e reduzindo a aleatoriedade dos dados
2. Suavização de Sazonalidade que combina a suavização dos dados com um ajustamento para o comportamento sazonal

Métodos de Previsão através de Modelos com Nenhuma Tendência ou Sazonalidade

A previsão é que os valores futuros serão constantes. Isto porque o modelo não possui tendência e assume-se que a oscilação de curto prazo é apenas **ruído**.

Aqui a série temporal possui aleatoriedade, mas não possui sazonalidade.

Média Móvel Simples (MMS)

O método da *média móvel simples* é indicado para previsões de curto prazo onde as componentes de *tendência* e *sazonalidade* são **inexistentes** ou possam ser desprezadas (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998).

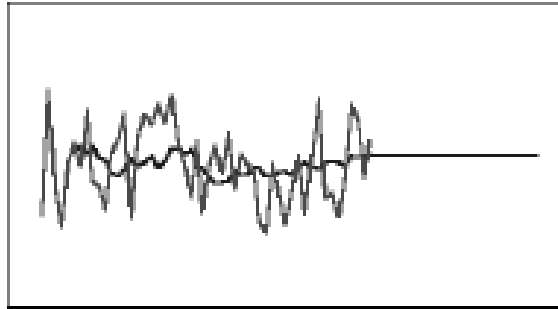


Figura 5 – Gráfico típico de dados de média móvel simples, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Esse modelo nada mais é do que uma técnica simples de previsão exponencial onde são considerados os **k** últimos dados históricos e, com estes, é realizada uma média aritmética (ou ponderada) para prever o valor do próximo dado. O número de observações *n* em cada cálculo da média, ou período, permanece constante e estipulado de maneira a tentar eliminar da melhor forma possível as componentes de *tendência* e *sazonalidade*. (CHAMBERS; MULLICK; SMITH, 1971; ARCHER, 1980; MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998).

Este método usa os *N* últimos valores da série temporal x_t , como previsão para o tempo $t + 1$. Portanto:

$$\hat{x}_{t+1} = \frac{1}{N} \sum_{i=t-N+1}^t x_i \quad 1$$

Neste método o parâmetro usado é o número de pontos ou período *N*.

Como escolher o *N*? Queremos o valor de *N* que nos dê a melhor acurácia de previsão (i.é., minimize o desvio absoluto médio, MAD, ou o erro quadrático médio, MSE, da previsão). Estes dois conceitos serão tratados no Apêndice A.

Vamos aqui olhar para um exemplo de produção de leite da *Estância Tio Fred*, cujos dados podem ser encontrados na planilha *Dados de Leite* da pasta *Séries Temporais*. A pasta encontra-se junto com o presente texto, no arquivo zipado, *MetodosBasicosDePrevisao.zip*.

	A	B	C	D	E	F
1	Exemplo de Média Móvel para Produção de Leite					
2	Usar N = 3					
3						
4		yi	yhati	 ei 	ei^2	
5	Dia	Litros	Previsto	erro absoluto	erro quadrático	
6	1	212				
7	2	220				
8	3	170				
9	4	189				
10	5	197				
11	6	208				
12	7	216				
13	8	201				
14	9	204				
15	10	182				
16	11	193				
17	12	204				
18	13	189				
19	14	182				
20	15	201				
21	16	204				
22	17	197				
23	18	208				
24	19	201				
25	20	185				
26						
27			MAD			MSE
28						
29			se			
30						
31						

Figura 6 Dados da produção diária de leite da Estância Tio Fred.

Usamos N = 3, e entramos com a fórmula =MÉDIA(B6:B8) na célula C9. Esta fórmula é simplesmente:

$$\frac{y_t + y_{t-1} + y_{t-2}}{3}$$

Vaja a Figura 7 abaixo:

	A	B	C	D	E	F	G	
1	Exemplo de Média Móvel para Produção de Leite							
2	Use N = 3							
3								
4		y_i	y_h^{ati}	 e_i 	e_i²			
5	Dia	Litros	Previsto	erro absoluto	erro quadrático			
6	1	212						
7	2	220						
8	3	170						
9	4	189	200,7	11,7	136,1			
10	5	197	193,0	4,0	16,0			
11	6	208	185,3	22,7	513,8			
12	7	216	198,0	18,0	324,0			
13	8	201	207,0	6,0	36,0			
14	9	204	208,3	4,3	18,8			
15	10	182	207,0	25,0	625,0			
16	11	193	195,7	2,7	7,1			
17	12	204	193,0	11,0	121,0			
18	13	189	193,0	4,0	16,0			
19	14	182	195,3	13,3	177,8			
20	15	201	191,7	9,3	87,1			
21	16	204	190,7	13,3	177,8			
22	17	197	195,7	1,3	1,8			
23	18	208	200,7	7,3	53,8			
24	19	201	203,0	2,0	4,0			
25	20	185	202,0	17,0	289,0			
26	21		198,0					
27			MAD	10,2	153,2	MSE		
28								
29			se	12,7				
30								
31								

Figura 7 Planilha mostrando a previsão de produção de leite da Estância tio Fred.

Depois de preenchido com a fórmula até C25. Tomamos o erro absoluto na coluna D e o erro quadrático da coluna E. Ver todas as fórmulas entradas na Figura 8 abaixo.

	A	B	C	D	E	F	G
1	Exemplo de Média Móvel para Produção de Leite						
2	Use N = 3						
3							
4		yi	yhati	ei	ei^2		
5	Dia	Litros	Previsto	erro absoluto	erro quadrático		
6	1	212					
7	2	220					
8	3	170					
9	4	189	=MÉDIA(B6:B8)	=ABS(B9-C9)	=D9^2		
10	5	197	=MÉDIA(B7:B9)	=ABS(B10-C10)	=D10^2		
11	6	208	=MÉDIA(B8:B10)	=ABS(B11-C11)	=D11^2		
12	7	216	=MÉDIA(B9:B11)	=ABS(B12-C12)	=D12^2		
13	8	201	=MÉDIA(B10:B12)	=ABS(B13-C13)	=D13^2		
14	9	204	=MÉDIA(B11:B13)	=ABS(B14-C14)	=D14^2		
15	10	182	=MÉDIA(B12:B14)	=ABS(B15-C15)	=D15^2		
16	11	193	=MÉDIA(B13:B15)	=ABS(B16-C16)	=D16^2		
17	12	204	=MÉDIA(B14:B16)	=ABS(B17-C17)	=D17^2		
18	13	189	=MÉDIA(B15:B17)	=ABS(B18-C18)	=D18^2		
19	14	182	=MÉDIA(B16:B18)	=ABS(B19-C19)	=D19^2		
20	15	201	=MÉDIA(B17:B19)	=ABS(B20-C20)	=D20^2		
21	16	204	=MÉDIA(B18:B20)	=ABS(B21-C21)	=D21^2		
22	17	197	=MÉDIA(B19:B21)	=ABS(B22-C22)	=D22^2		
23	18	208	=MÉDIA(B20:B22)	=ABS(B23-C23)	=D23^2		
24	19	201	=MÉDIA(B21:B23)	=ABS(B24-C24)	=D24^2		
25	20	185	=MÉDIA(B22:B24)	=ABS(B25-C25)	=D25^2		
26	21		=MÉDIA(B23:B25)				
27			MAD	=MÉDIA(D9:D25)	=MÉDIA(E9:E25)	MSE	
28							
29			se	=1,25*D27			
30							
31							
32							
33							
34							
35							

Figura 8 Fórmulas entradas nas células da planilha Dados de Leite.

As desvantagens desse modelo estão relacionadas à falta de acurácia ao lidar com séries históricas que apresentam **tendência** ou **sazonalidade** já que, nesse método, a previsão para o próximo período envolve sempre a adição de novos dados e a desconsideração dos anteriores. Uma alternativa para amenizar esse erro é a utilização da média ponderada para tentar construir um padrão mais próximo à realidade. A dificuldade na utilização da média móvel ponderada é a necessidade de conhecimento para se determinar os pesos a serem utilizados (DAVIS; AQUILANO; CHASE, 2001). Mesmo assim, a título de ilustração, vamos abordar esse assunto a seguir.

Média Móvel Ponderada (WMA)

No método da média móvel acima, cada observação na média é ponderada igualmente por isso o método de média móvel ponderada, tipicamente as observações mais recentes carregam pesos maiores na média. A expressão para a média móvel ponderada é:

$$\hat{y}_t = w_t * y_t + w_{t-1} * y_{t-1} + \dots + w_{t-N+1} * y_{t-N+1}$$

Seja w_i = o peso para a observação i

$\sum w_i = 1 \Rightarrow$ a soma dos pesos é 1.

Para ilustrar, vamos reescrever o exemplo do leite. Vá à planilha *Solução WMA* que se encontra na pasta *Séries Temporais*. Usar um peso de 0,5 e entrar com ele na célula H2 para a observação mais recente, um peso de 0,3, na célula H3 para a próxima mais recente observação, e um peso de 0,2 na célula H4 para a observação mais antiga (ver Figura 9).

	A	B	C	D	E	F	G	H	I	
1	Exemplo de Média Móvel Ponderada para a Produção de Leite									
2	Usar N = 3							w(l-1)	0,50	
3							w(l-2)	0,30		
4		yi	yhati	ei	ei*2		w(l-3)	0,20		
5	Dia	Litros	Previsto	erro absoluto	erro quadrático			1,00		
6	1	212								
7	2	220								
8	3	170								
9	4	189	193,4	4,4	19,4					
10	5	197	189,5	7,5	56,3					
11	6	208	189,2	18,8	353,4					
12	7	216	200,9	15,1	228,0					
13	8	201	209,8	8,8	77,4					
14	9	204	206,9	2,9	8,4					
15	10	182	205,5	23,5	552,3					
16	11	193	192,4	0,6	0,4					
17	12	204	191,9	12,1	146,4					
18	13	189	196,3	7,3	53,3					
19	14	182	194,3	12,3	151,3					
20	15	201	188,5	12,5	156,3					
21	16	204	192,9	11,1	123,2					
22	17	197	198,7	1,7	2,9					
23	18	208	199,9	8,1	65,6					
24	19	201	203,9	2,9	8,4					
25	20	185	202,3	17,3	299,3					
26			194,4							
27			MAD	9,8	135,4	MSE				
28										
29			se	12,3						
30										
31										

Figura 9 Planilha de cálculo para o método da Média Móvel Ponderada (WMA)

Novamente usando $N = 3$, entre com a fórmula $= (B8*\$H\$2 + B7*\$H\$3 + B6*\$H\$4)$ na célula C9 e preencha até a C25. (Nota: usamos o \$ para bloquear a célula em H2, H3 e H4).

O erro absoluto e o erro quadrático são entrados na coluna D e E, respectivamente. A célula D27 mostra o MAD e a célula E27 é o MSE.

Poderíamos ter usado um conjunto melhor de pesos que aquele usado acima para ter uma previsão melhor? Vamos usar o Solver do Excel para verificar isto para você. Na guia Dados, no grupo Análise, clique em Solver. Assim

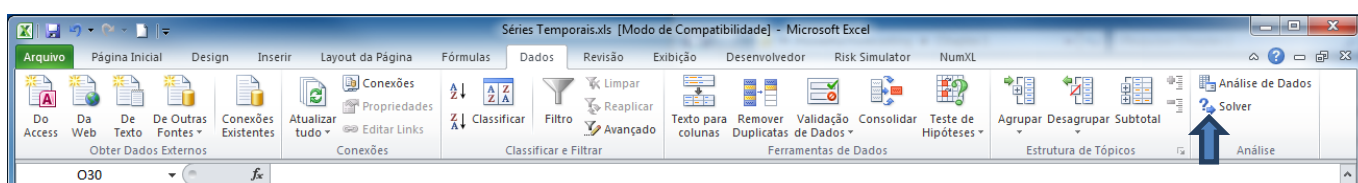


Figura 10 A faixa de opções (ribbon) do Excel mostrando a ferramenta Solver

Uma caixa de diálogo aparecerá como na Figura 11 abaixo

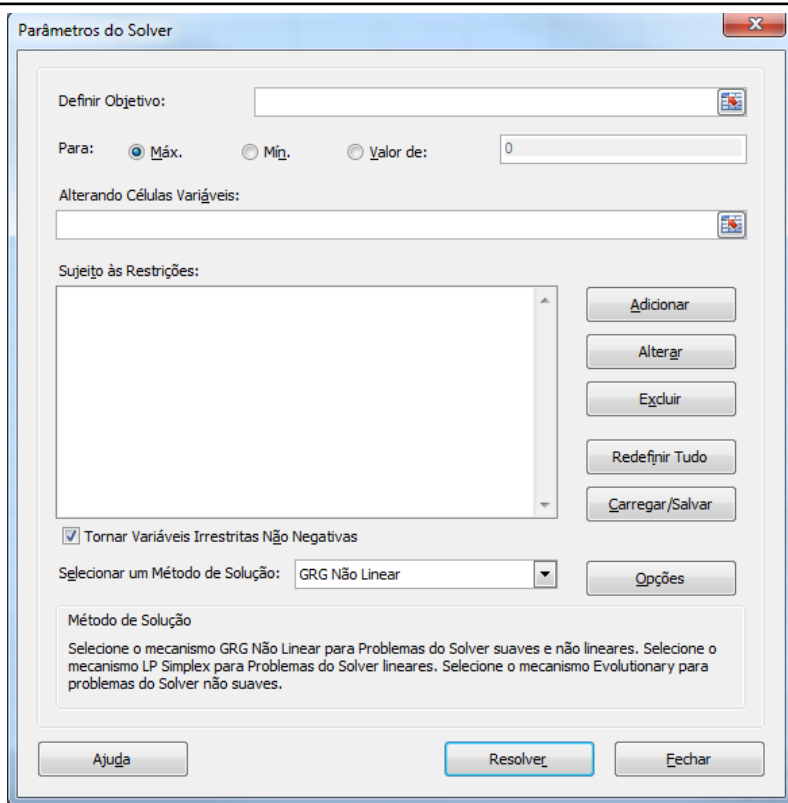


Figura 11 – A caixa de diálogo Parâmetros do Solver.

A seguir entramos com todos os parâmetros na caixa de diálogo. Ela se parecerá como a Figura 12 abaixo após todos os parâmetros terem sido entrados.

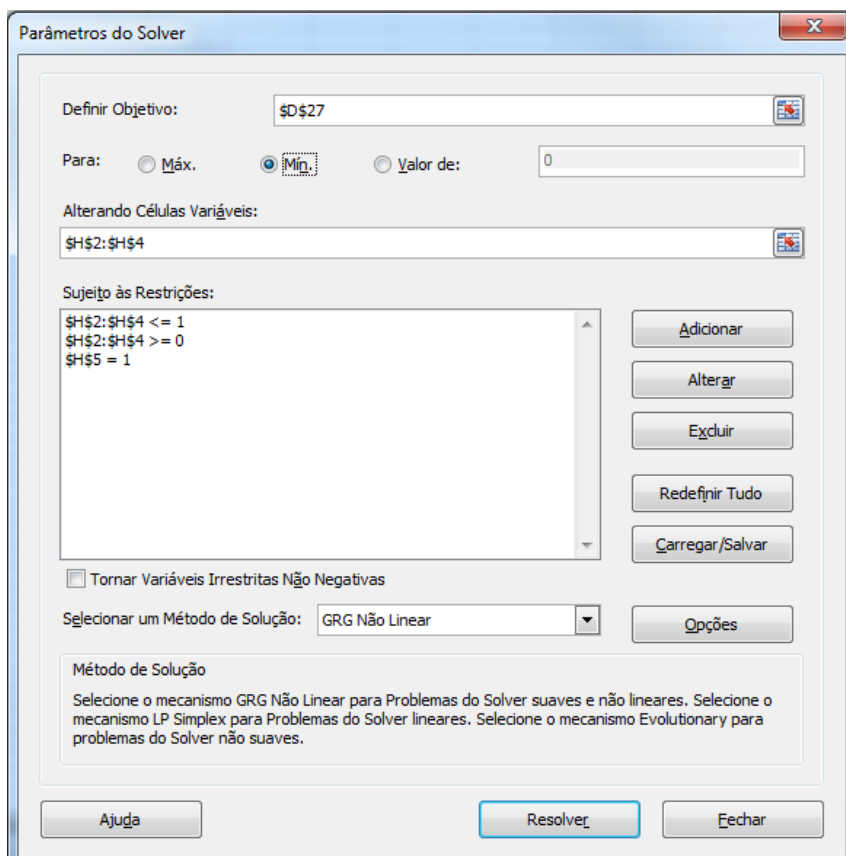


Figura 12 – Após os parâmetros terem sido introduzidos na caixa de diálogo.

Clique no botão Resolver. O Solver iniciará a otimização.

No final o Solver lança a janela:

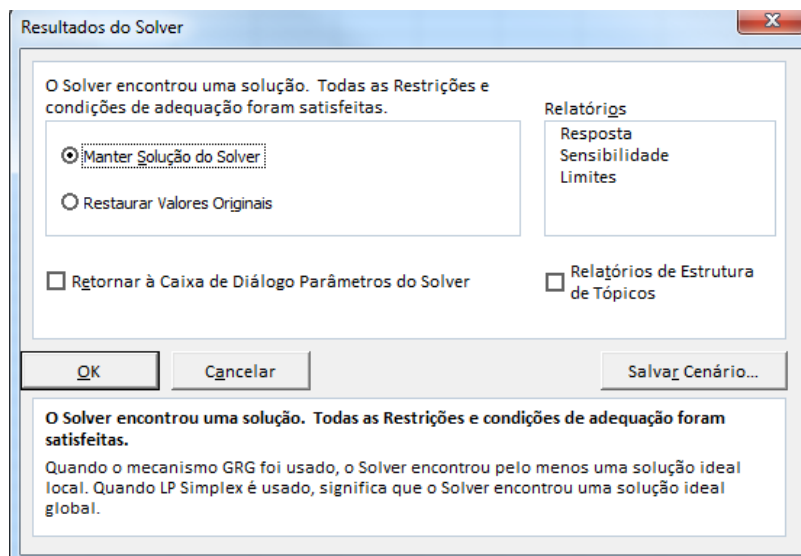


Figura 13 – Caixa de mensagem Resultados do Solver

Inicialmente tínhamos o MAD = 9,8 na célula D27 (ver Figura 9 acima). Após a otimização com o Solver do Excel o MAD agora é igual a 8,7 e os pesos são 0,57,0,11,0,32 no intervalo de células H2:H4, respectivamente.

	A	B	C	D	E	F	G	H	I	
1	Exemplo de Média Móvel Ponderada para a Produção de Leite									
2	Usar N = 3						w(l-1)	0,57		
3							w(l-2)	0,11		
4		yi	yhati	 eij	ei^2	w(l-3)		0,32		
5	Dia	Litros	Previsto	erro absoluto	erro quadrático			1,00		
6	1	212								
7	2	220								
8	3	170								
9	4	189	189,0	0,0	0,0					
10	5	197	197,0	0,0	0,0					
11	6	208	187,4	20,6	424,9					
12	7	216	200,7	15,3	235,4					
13	8	201	209,0	8,0	63,7					
14	9	204	204,9	0,9	0,8					
15	10	182	207,6	25,6	653,6					
16	11	193	190,5	2,5	6,1					
17	12	204	195,4	8,6	74,3					
18	13	189	195,7	6,7	44,7					
19	14	182	191,9	9,9	98,3					
20	15	201	189,9	11,1	123,6					
21	16	204	195,1	8,9	79,9					
22	17	197	196,5	0,5	0,2					
23	18	208	199,1	8,9	80,1					
24	19	201	205,5	4,5	20,4					
25	20	185	200,5	15,5	238,9					
26			194,2							
27			MAD	8,7	126,2	MSE				
28										
29			se	10,8						
30										

Figura 14 – Resultados encontrados pelo Solver.

Como você pode ver com este exemplo, o método da média móvel ponderada é mais complicado para se construir e dará um resultado melhor.

Uma das desvantagens de se usar médias móveis para previsão é que no cálculo à média de todas as observações são dados pesos iguais (ou seja $1/L$), enquanto esperaríamos que as observações mais recentes sejam um melhor indicador do futuro (e assim deveria ser dado maior peso). Também, nas médias móveis usamos somente observações recentes, talvez devêssemos considerar todas as observações anteriores.

Uma técnica (que olharemos para ela a seguir) conhecida como suavização exponencial (ou, mais precisamente, suavização exponencial simples) dá peso maior às observações mais recentes e considera todas as observações anteriores.

Suavização Exponencial Simples

Muito usado nos dias de hoje é o modelo de *suavização exponencial simples* por ser extremamente simples e possuir fácil capacidade de ajustes em relação à acurácia obtida com esse método. Enquanto na média móvel simples as observações passadas são ponderadas igualmente, a suavização exponencial atribui *pesos decrescendo exponencialmente* quando a observação ficar mais velha. Em outras palavras, *observações recentes são dadas relativamente mais peso na previsão do que as observações mais antigas*.

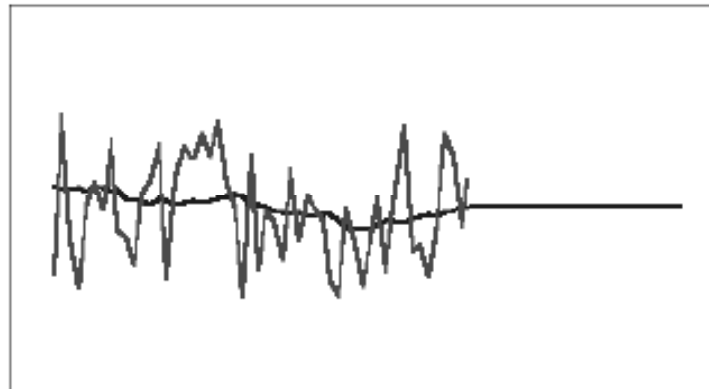


Figura 15 – Gráfico típico de dados de suavização exponencial simples, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Caso a série temporal em estudo mantenha-se constante sobre um nível médio, uma suavização exponencial simples pode ser usada para a previsão de valores futuros dessa série. A representação matemática desse modelo é dada por (MAKRIDAKIS et al., 1998)

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t \quad 2$$

onde \hat{y}_{t+1} é a previsão da demanda para o tempo $t+1$, feita no período atual t ; α é a **constante de suavização**, assumindo valores entre 0 e 1; y_t é o valor observado (real) na série temporal para o tempo t ; e, \hat{y}_t é o valor da previsão feita para o tempo t .

Uma forma de medir a acurácia da previsão, é calculando o erro gerado pela mesma, ou seja:

$$\varepsilon_t = y_t - \hat{y}_t \quad 3$$

O valor da constante de suavização α é arbitrário². Pode-se determinar o melhor valor para esta através de métodos iterativos para minimizar alguma medida de qualidade da previsão como, por exemplo, a média do quadrado dos erros, EQM ou o a sua raiz quadrada, RMSE. Desta maneira, seleciona-se, inicialmente,

² O CB Predictor e o Risk Simulator calculam automaticamente a constante de suavização ótima.

um valor aleatório para a constante, a partir do qual previsões são geradas. Comparam-se os valores previstos com os reais, e calcula-se a média do quadrado das diferenças entre os mesmos; o parâmetro que minimiza essa média é utilizado no modelo final³. A magnitude da constante α determina a velocidade de resposta do modelo frente a mudanças valores da série (MONTGOMERY et al., 1990). Valores baixos para a constante α faz com que o modelo demore a reagir às mudanças no comportamento da série. Com valores altos de α , o modelo responde rapidamente.

Os modelos de suavização exponencial simples requerem uma estimativa inicial para \hat{x}_t . Quando dados históricos estão disponíveis, pode-se usar uma média simples das N observações mais recentes como \hat{x}_t ; caso contrário, pode-se utilizar a observação mais recente, ou fazer uma estimativa subjetiva.

Uma medida de eficiência deste método pode ser obtida sob a consideração que o processo é completamente estável, assim que X_1, X_2, \dots , são variáveis aleatórias independentes e identicamente distribuído (IID)⁴ com variância σ^2 . Portanto, segue que (para grande t):

$$\text{var}[\hat{x}_{t+1}] \approx \frac{\alpha\sigma^2}{2-\alpha} = \frac{\sigma^2}{(2-\alpha)/\alpha} \quad 4$$

Assim que a variância é estatisticamente equivalente para a média móvel com $(2 - \alpha)/\alpha = 19$. Assim, em termos de variância, o Método de Suavização Exponencial com este valor de α é equivalente ao Método da Média Móvel que utiliza 19 observações. Entretanto, se uma mudança no processo ocorre, a Suavização Exponencial irá reagir mais rapidamente com melhor ajuste de que o Método da Média Móvel.

Uma desvantagem deste Método está na dificuldade em escolher um valor apropriado para α . O Método de Suavização Exponencial pode ser visto como um processo de filtragem com um filtro estatístico cujas entradas são os dados “puros” a partir de um processo estocástico e a saída são estimativas suavizadas de uma média que varia com o tempo.

Uma maneira de iniciar o processo *recursivo* é utilizar $\hat{y}_1 = 0$ e $\hat{y}_2 = y_1$.

Vamos agora reescrever o problema do LEITE. Vá à planilha *Solução EXP*. Use uma suavização exponencial simples com $\alpha = 0,3$ como valor inicial na célula G4.

³ Pode-se lançar mão do Solver do Excel para realizar isto.

⁴ Duas variáveis aleatórias são independentes se $P(A \cap B) = P(A) \cdot P(B) = P(A) \cdot P(B)$.

Duas variáveis aleatórias são identicamente distribuídas se possuem a mesma distribuição de probabilidade.

	A	B	C	D	E	F	G	H
1	Exemplo de Suavização Exponencial Simples para a Produção de Leite							
2	Usar alfa = 0,3							
3							alfa	
4		yi	yhati	 ei 	ei^2		0,3	
5	Dia	Litros	Previsto	erro absoluto	erro quadrático			
6	1	212	212,0					
7	2	220	212,0	8,0	64,0			
8	3	170	214,4	44,4	1971,4			
9	4	189	201,1	12,1	145,9			
10	5	197	197,5	0,5	0,2			
11	6	208	197,3	10,7	114,1			
12	7	216	200,5	15,5	239,5			
13	8	201	205,2	4,2	17,4			
14	9	204	203,9	0,1	0,0			
15	10	182	203,9	21,9	481,4			
16	11	193	197,4	4,4	19,0			
17	12	204	196,1	7,9	63,2			
18	13	189	198,4	9,4	89,0			
19	14	182	195,6	13,6	185,1			
20	15	201	191,5	9,5	89,8			
21	16	204	194,4	9,6	92,8			
22	17	197	197,3	0,3	0,1			
23	18	208	197,2	10,8	117,1			
24	19	201	200,4	0,6	0,3			
25	20	185	200,6	15,6	243,3			
26			195,9					
27			MAD	10,5	207,0	MSE		
28								
29			se	13,1				
30								

Figura 16 A suavização exponencial simples aplicada à produção de leite da Estância Tio Fred.

É costume assumir $\hat{y}_{t_1} = y_1$ por isso entramos com C6 = B6. Após isto inicie o período 2 na célula e C7 entramos com a fórmula $=\$G\$4*B6+(1-\$G\$4)*C6$ e preencha com a fórmula até a célula C25. (Nota: usamos o símbolo \$ para travar a célula G4). Desse jeito é como se entra com a fórmula $\hat{y}_{t+1} = \hat{y}_t + \alpha (y_t - \hat{y}_t)$ no Excel.

O erro absoluto e o erro quadrático são entrados nas colunas D e E, respectivamente. A célula D27 mostra o MAD e a célula E27 é o MSE. Usaremos o Solver do Excel para encontrar o melhor valor para α e também minimizaremos o MAD e MSE.

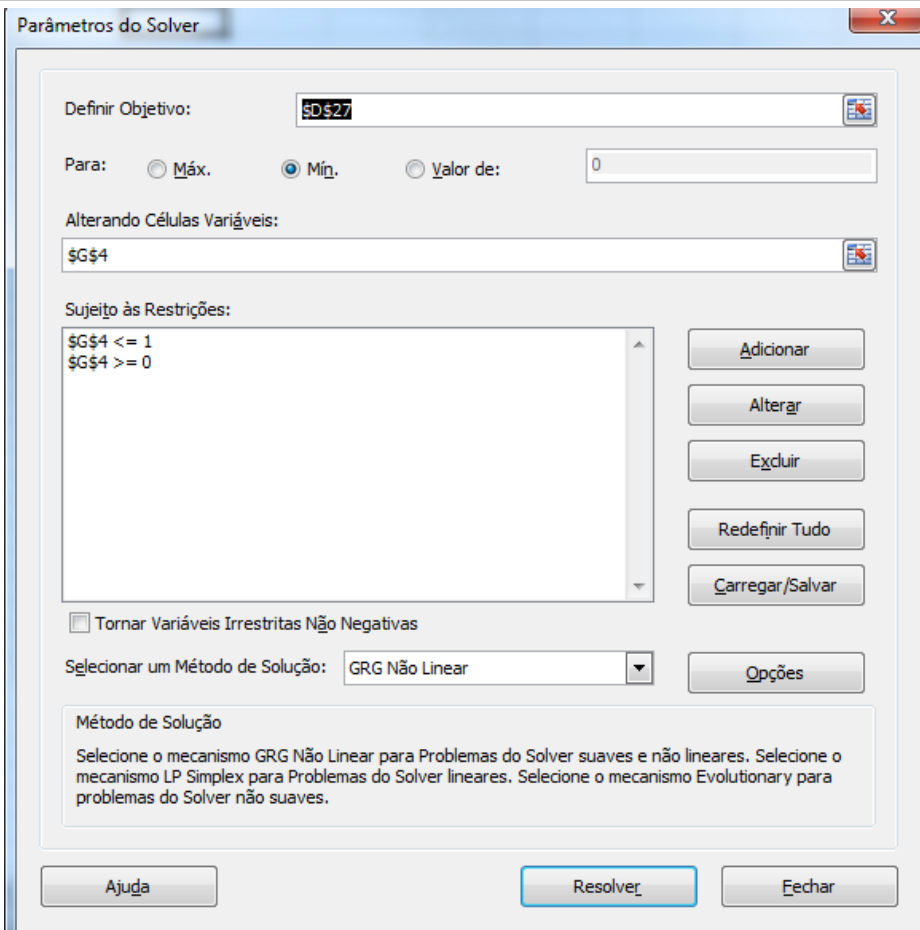


Figura 17 – A caixa de diálogo do Solver com os parâmetros para a suavização exponencial simples. Entramos com todos os parâmetros na caixa de diálogo. Ela se parecerá com a Figura 17 acima após todos os parâmetros serem entrados.

Definir objetivo: MAD(D27); Para: Min

Alterando Células Variáveis: o peso em G4

Sujeito às Restrições: $0 \leq G4 \leq 1$

Clique no botão Resolver. O Solver iniciará a otimização.

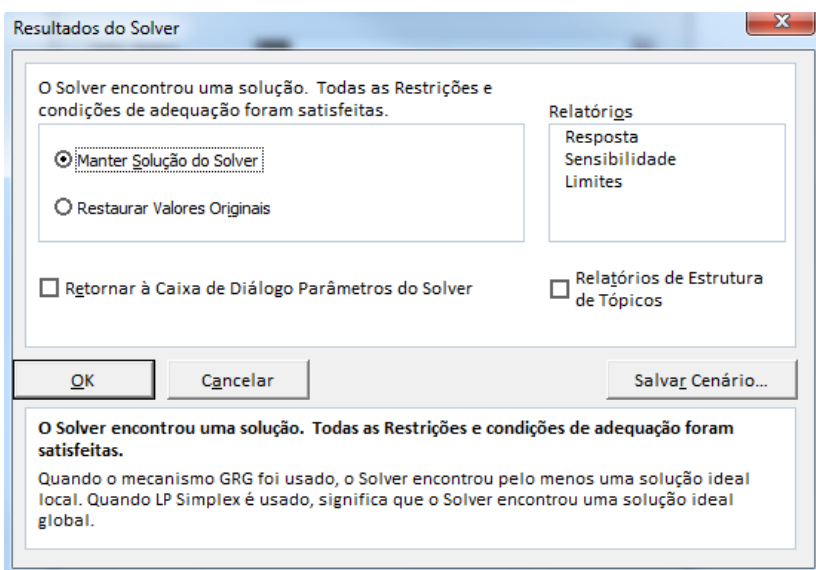


Figura 18 – A janela de Resultados do Solver

Clique em OK e mantenha a solução do Solver.

O MAD na célula D27 é igual a 10,5 e o α na célula G4 é igual a 0,311737. Melhoramos um pouco.

Se no dia 20 tivermos 185 litros, como você fará a previsão da produção no dia 21? Para prever simplesmente preencha a fórmula da célula 25 na célula 26 ou entre com a fórmula $=\$G\$4*B25+(1-\$G\$4)*C25$. Aqui está o resultado: 195,8 litros (ver Figura 19).

	A	B	C	D	E	F	G	H
1	Exemplo de Suavização Exponencial Simples para a Produção de Leite							
2	Usar alfa = 0,3							
3							alfa	
4		yi	yhati	 lei 	ei^2		0,311737	
5	Dia	Litros	Previsto	erro absoluto	erro quadrático			
6	1	212	212,0					
7	2	220	212,0	8,0	64,0			
8	3	170	214,5	44,5	1979,7			
9	4	189	200,6	11,6	135,1			
10	5	197	197,0	0,0	0,0			
11	6	208	197,0	11,0	121,0			
12	7	216	200,4	15,6	242,5			
13	8	201	205,3	4,3	18,3			
14	9	204	203,9	0,1	0,0			
15	10	182	204,0	22,0	482,4			
16	11	193	197,1	4,1	17,0			
17	12	204	195,8	8,2	66,7			
18	13	189	198,4	9,4	88,0			
19	14	182	195,5	13,5	181,1			
20	15	201	191,3	9,7	94,8			
21	16	204	194,3	9,7	94,1			
22	17	197	197,3	0,3	0,1			
23	18	208	197,2	10,8	116,2			
24	19	201	200,6	0,4	0,2			
25	20	185	200,7	15,7	246,9			
26			195,8					
27			MAD	10,5	207,8	MSE		
28								
29			se	13,1				
30								

Figura 19 Resultados da previsão de produção de leite da Estância Tio Fred, usando o Solver para minimizar o erro MAD.

Como você pode ver neste exemplo, o método de suavização exponencial simples é um pouco mais complicado para se construir e dá-nos um resultado melhor.

A suavização exponencial é útil quando não há tendência. Entretanto, se os dados tiverem tendência, precisamos usar o método de Suavização Exponencial Dupla que será discutido abaixo.

Métodos de Previsão para Modelos com Tendência e Nenhuma Sazonalidade

A representação matemática para o processo (que gera a série temporal) com valor constante, *tendência* e *flutuações aleatórias* pode ser dada por:

$$X_t = a t + \pi + \varepsilon_t$$

5

Com $t = 1, 2, \dots$ e onde:

X_t é uma variável aleatória observada no tempo t ;

a é a tendência do modelo;

π é o valor constante do modelo;

ε_t é o erro aleatório ocorrido no tempo t (geralmente assumido ter valor esperado igual a zero e variância constante).

Modelo de Previsão com Média Móvel Dupla

Aqui a técnica de média móvel simples é aplicada duas vezes, uma nos dados originais e depois nos dados resultantes desta primeira aplicação.

Define-se a média móvel dupla de tamanho k como:

$$M_T^{[2]} = \frac{M_T^{[1]} + M_{T-1}^{[1]} + \dots + M_{T-p+1}^{[1]}}{p} \quad 6$$

Onde $M_T^{[1]}$ é a média móvel (simples) de tamanho p calculada usando todas as observações até o instante T (inclusive).

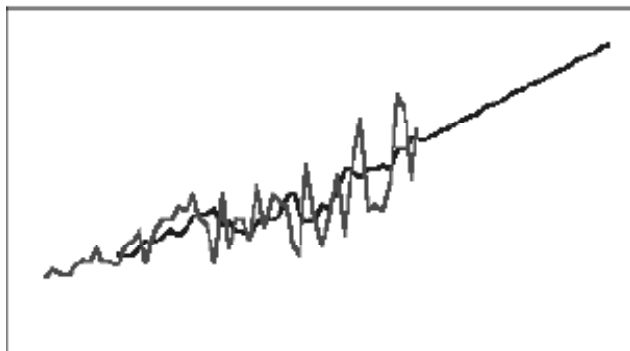


Figura 20 – Gráfico típico de dados de média móvel dupla, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Por que usar médias móveis duplas?

Se os dados exibem uma tendência linear, o uso de médias móveis simples para a previsão dos valores da série induz a erros sistemáticos, pois a média móvel simples segue a tendência com certo atraso, e este efeito é amplificado quando tentamos prever valores futuros. O método de médias móveis duplas procura diminuir este efeito sistemático.

A previsão é realizada por meio de uma reta inclinada, isto é, há uma expectativa de que o valor da variável será sempre crescente de modo a ser compatível com a tendência nos dados históricos.

Os dados possuem aleatoriedade e tendência de crescimento, mas veja que não há sazonalidade. Atribui a todos os valores passados o mesmo peso na previsão. Veja que em previsões o analista deve se preocupar com a tendência. Por quê?

O modelo matemático para a previsão de k períodos com média móvel é

$$\hat{x}_{t+k} = 2 * M_t^{[1]} - M_t^{[2]} + \frac{2p}{p-1} (M_t^{[1]} - M_t^{[2]}) \quad 7$$

onde:

$M_t^{[1]}$ é a média móvel(simples) de tamanho k calculada usando as k observações anteriores ao instante t (inclusive);

$M_t^{[2]}$ é a média móvel(dupla) de tamanho k calculada usando as k médias móveis simples, $M_t^{[1]}$ anteriores ao instante t (inclusive).

p é o período usado no cálculo da média móvel;

k é o número de períodos de previsão variando de 1 até h (horizonte de previsão).

Modelo de Previsão por Alisamento Exponencial Duplo de Holt⁵

Quando uma determinada série apresenta aleatoriedade e uma tendência *linear* de crescimento (ou decrescimento), o modelo de suavização exponencial dupla de **Holt** pode ser usado de maneira satisfatória para a previsão, caso os outros componentes da série possam ser desprezados. Este modelo emprega duas constantes de suavização, α e β (com valores entre 0 e 1), sendo representado por três equações (MAKRIDAKIS et al., 1998):

$$L_t = \alpha x_t + (1 - \alpha)(E_{t-1} + T_{t-1}) \quad 8$$

$$T_t = \beta(L_t - E_{t-1}) + (1 - \beta)T_{t-1} \quad 9$$

$$\hat{y}_{t+k} = E_t + kT_t \quad 10$$

onde:

E_t é a componente de nível;

T_t é a componente de tendência;

h é o horizonte de previsão;

$k = 1, 2, \dots, h$;

\hat{y}_{t+k} é a previsão;

α com valores no intervalo $0 < \alpha < 1$, é a constante de suavização da componente de **nível** L_t ;

β com valores no intervalo $0 < \beta < 1$, é a constante de suavização da componente **tendência** T_t ;

Na equação 8, pode-se perceber que o valor de nível E_t é a média ponderada do próprio valor da série y_t e de E_{t-1} e T_{t-1} (nível e tendência previstos no tempo $t-1$, respectivamente). Para uma série isenta de erro aleatório, a quantidade $(E_{t-1} + T_{t-1})$ é exatamente o valor de E_t , uma vez que a variação de tempo entre t e $t-1$ é obviamente 1. Assim, a expressão (8) pode ser entendida como:

$$E_t = \alpha y_t + (1 - \alpha)(E_t) \quad 11$$

onde:

$$E_t = f(L_{t-1}, T_{t-1}) \quad 12$$

Em (9), a parcela $E_t - E_{t-1}$ é a derivada discreta que representa, portanto, a tendência. Para o restante, o raciocínio é análogo ao realizado para a expressão (8).

Considerando que a primeira amostra da série temporal é para $t = 1$, os valores L_1 e T_1 são funções de L_0 e T_0 . Como não existe amostra da série para $t = 0$, faz-se necessário inicializar L_1 e T_1 . Há várias maneiras de se inicializar estas variáveis, dentre as quais:

$$E_1 = y_1 \quad 13$$

$$T_1 = y_2 - y_1 \quad 14$$

⁵ Muito usado em Engenharia de Confiabilidade

ou

$$T_1 = \frac{y_t - y_1}{t - 1} \quad 15$$

$$T_1 = \frac{(y_2 - y_1) + (y_3 - y_2) + (y_4 - y_3)}{3} \quad 16$$

OBS:- Uma vez que a componente de tendência em uma série é representada apenas por um coeficiente (coeficiente angular da reta) as formas apresentadas em (14), (15), (16) para inicializar T_1 são possíveis representações para a derivada discreta da série calculada em $t = 1$.

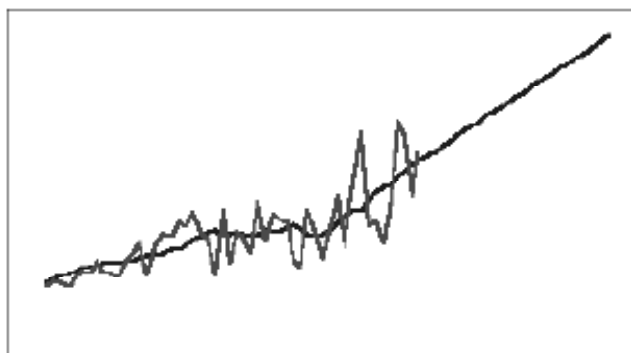


Figura 21 – Gráfico típico de dados de suavização exponencial dupla de Holt, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

O modelo é muito usado para modelagem de produtos na fase de divulgação quando começa a ser aceito pelo público consumidor.

As equações (8) e (9) fazem uma estimativa do *nível* e da *inclinação* da série temporal, respectivamente. Já a equação (10), calcula a previsão da série para os próximos k períodos.

Assim como na suavização exponencial simples, o método de Holt requer valores iniciais, neste caso, E_0 e T_0 . Uma alternativa para estes cálculos iniciais é igualar E_0 ao último valor observado na série temporal e calcular uma média da declividade nas últimas observações para T_0 . Outra forma de cálculo é a regressão linear simples aplicada aos dados da série temporal, onde se obtém o valor da declividade da série temporal e de E_0 em sua origem.

As constantes de suavização α e β no modelo de Holt podem ser determinadas de maneira análoga à obtenção de α na suavização exponencial simples, ou seja, através da utilização de um método iterativo que encontre a combinação de α e β que minimize o EQM.

Abra a planilha *Solução EXP DUPLA*.

Os dados são para as vendas mensais em milhares para uma empresa de vestuário. Inicialmente use o valor de 0,2 (célula I4) para α e 0,3 (célula J4) para β . Para $E_1 = y_1$, entramos com =B6 na célula C6.

Depois então entre com a fórmula = $\$I\$4*B7+(1-\$I\$4)*(C6+D6)$ na célula C7 e preencha até a C29. Como você pode ver na fórmula, iniciamos no período 2 na célula C7. (Nota: usamos o sinal \$ para travar a célula em I4). Isto é como a fórmula $E_t = \alpha y_t + (1 - \alpha) (E_{t-1} + T_{t-1})$ é entrada como fórmula no Excel.

Depois disto, calcularemos $T_t = \beta (E_t - E_{t-1}) + (1 - \beta) T_{t-1}$. É prática comum entrar com 0 no período 1, i.é., célula D6. Iniciando no período 2, na célula D7, entramos com a fórmula = $\$J\$4*(C7-C6)+(1-\$J\$4)*D6$ e preenchemos abaixo até D29. Isto é como a fórmula $T_t = \beta (E_t - E_{t-1}) + (1 - \beta) T_{t-1}$ é entrada como fórmula no Excel.

Agora junte as duas fórmulas entrando com =C6+D6 em E7 e preencha abaixo até E29, i.é., a fórmula

$\hat{y}_{t+k} = E_t + kT_t$. O erro absoluto e o erro quadrático são entrados nas colunas F e G, respectivamente. A célula F31 mostra o MAD e a célula G31 é o MSE. Usaremos o Solver do Excel para minimizar o MAD e encontrar o melhor valor ou otimizar α e β (ver Figura 22 abaixo).

	A	B	C	D	E	F	G	H	I	J
1	Exemplo de Suavização Exponencial com Tendência para Vendas de Roupas de Inverno									
2	Use alfa = 0,2 e beta = 0,3									
3									alfa	beta
4		yi	Ei	Ti	yhati	ei	ei^2		0,2	0,3
5	Meses	unidades	Base	Tendência	Previsto	erro absoluto	erro quadrático			
6	1	684	684	0						
7	2	590	665,20	-5,64	684,00	94,00	8836,00			
8	3	750	677,65	-0,21	659,56	90,44	8179,39			
9	4	880	717,95	11,94	677,43	202,57	41032,82			
10	5	885	760,91	21,25	729,89	155,11	24059,78			
11	6	788	783,33	21,60	782,16	5,84	34,14			
12	7	1004	844,74	33,54	804,92	199,08	39631,45			
13	8	1111	924,82	47,51	878,28	232,72	54158,13			
14	9	1160	1009,86	58,77	972,33	187,67	35219,97			
15	10	1044	1063,70	57,29	1068,63	24,63	606,62			
16	11	1500	1196,79	80,03	1120,99	379,01	143647,44			
17	12	1610	1343,46	100,02	1276,82	333,18	111007,93			
18	13	1250	1404,78	88,41	1443,48	193,48	37433,03			
19	14	1730	1540,55	102,62	1493,19	236,81	56078,34			
20	15	1990	1712,54	123,43	1643,17	346,83	120289,66			
21	16	2030	1874,77	135,07	1835,97	194,03	37649,11			
22	17	2100	2027,87	140,48	2009,84	90,16	8128,17			
23	18	1760	2086,68	115,98	2168,35	408,35	166753,74			
24	19	2300	2222,13	121,82	2202,66	97,34	9474,56			
25	20	2620	2399,16	138,38	2343,95	276,05	76204,10			
26	21	2566	2543,23	140,09	2537,54	28,46	809,90			
27	22	2710	2688,66	141,69	2683,32	26,68	711,68			
28	23	2800	2824,28	139,87	2830,35	30,35	921,02			
29	24	2850	2941,32	133,02	2964,15	114,15	13029,74			
30	25				3074,34					
31					MAD	171,61	43212,90	MSE		
32										
33					se	214,51				
34										
35										
36										
37	32		Previsão para o Mês 32		4005,482					
38										
39										

Figura 22 o Método de suavização exponencial dupla de Holt para uma empresa de vestuário

Invoque o Solver do Excel para minimizar o MAD. Entramos com todos os parâmetros na caixa de diálogo. Ela se parecerá como a Figura 23 abaixo, depois que todos os parâmetros forem entrados:

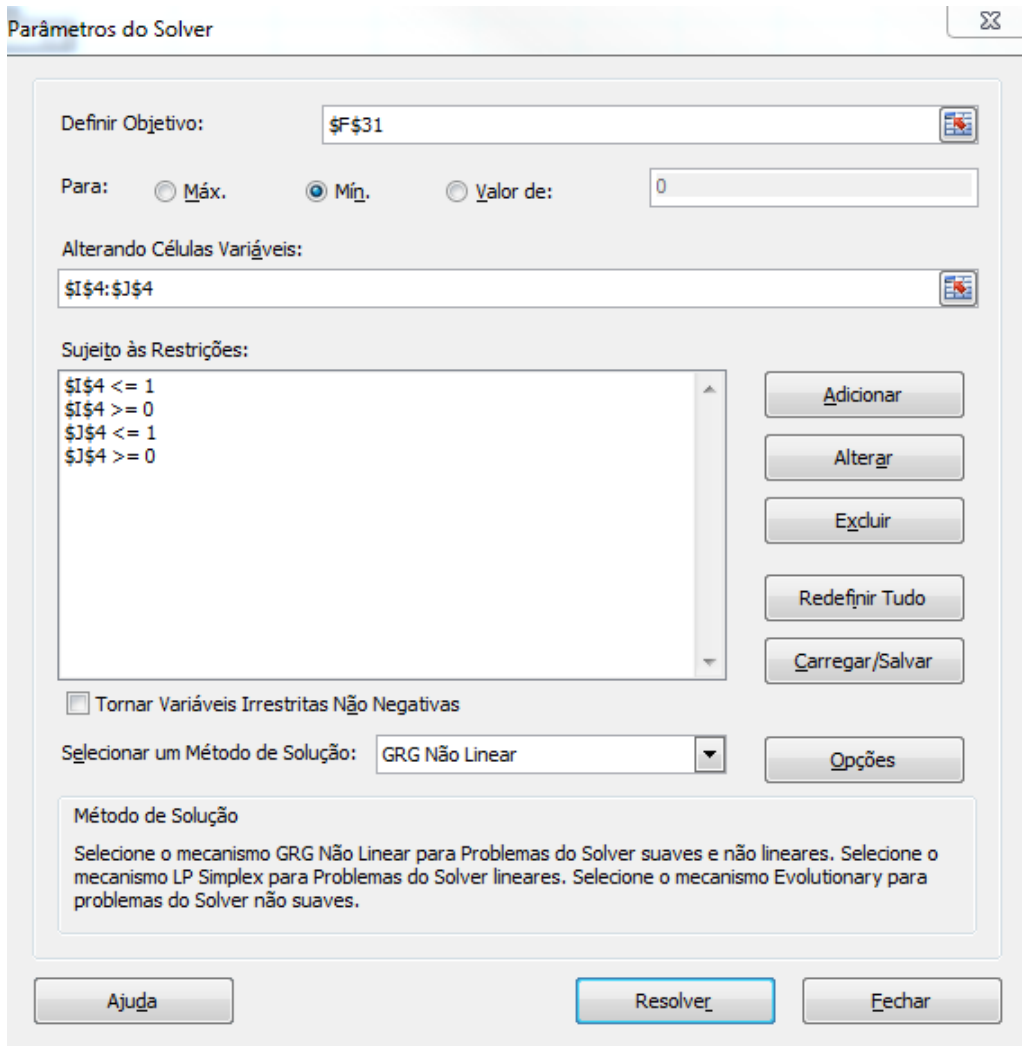


Figura 23 A caixa de diálogo do Solver depois de entrar os parâmetros

Definir objetivo: MAD(F31); Para: Mín

Alterando Células Variáveis: os pesos em I4:J4

Sujeito às Restrições: $0 \leq I4 \leq 1$, $0 \leq J4 \leq 1$

Clique no botão Resolver. O Solver iniciará a otimização.

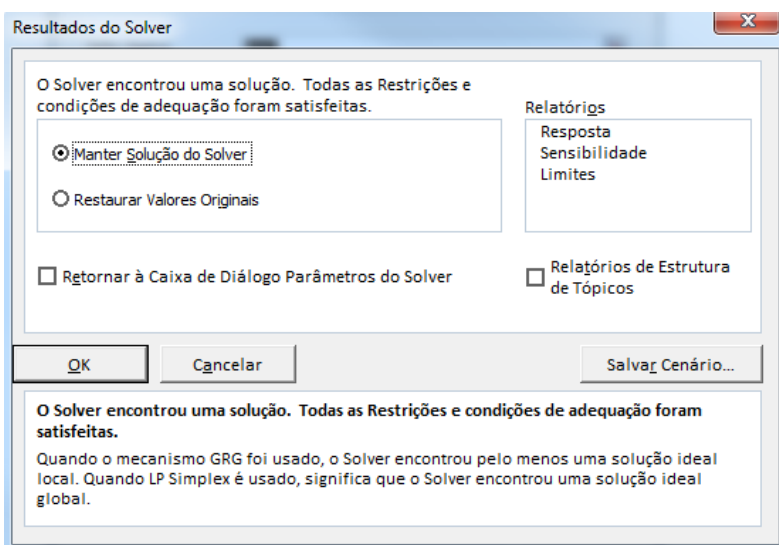


Figura 24 A janela de Resultados do Solver

Mantenha a solução do Solver.

O MAD na célula F31 foi minimizado e ficou igual a 140,91 e o α na célula I4 ficou igual a 0,126401. O β na célula J4 ficou igual a 1. Fizemos uma pequena melhora após a otimização com o Solver.

Se no mês 24 tivermos 2.850.000, como você faria para prever a produção no mês 25? Para prever simplesmente preencha a fórmula da célula E29 na célula E30 ou entre com a fórmula =C29+D29. Aqui o resultado é 3.037.810 (ver Figura 1.17 abaixo).

Se você quiser prever k períodos adiante, use a seguinte lógica

$$\hat{y}_{t+k} = E_t + k T_t$$

Neste exemplo, queremos prever as vendas no mês 32, i.é., 8 meses adiante, então entramos com =C29+8*\$D\$29 como você pode ver na célula E37. O resultado é 3.868.383 (ver Figura 1.17)

$E_t = C29$ $T_t = D29$ e $k = 8$

	A	B	C	D	E	F	G	H	I	J
1	Exemplo de Suavização Exponencial com Tendência para Vendas de Roupas de Inverno									
2	Use alfa = 0,2 e beta = 0,3									
3									alfa	beta
4									0,126401	1
5	Meses	yi	Ei	Ti	yhati	eij	ei^2			
	unidades	Base	Tendência	Previsto	erro absoluto	erro quadrático				
6	1	684	684	0						
7	2	590	672,12	-11,88	684,00	94,00	8836,00			
8	3	750	671,58	-0,54	660,24	89,76	8057,46			
9	4	880	697,46	25,88	671,05	208,95	43661,24			
10	5	885	743,77	46,31	723,34	161,66	26135,44			
11	6	788	789,82	46,05	790,08	2,08	4,33			
12	7	1004	857,12	67,30	835,87	168,13	28269,17			
13	8	1111	948,00	90,88	924,42	186,58	34812,79			
14	9	1160	1054,20	106,19	1038,89	121,11	14668,47			
15	10	1044	1145,68	91,48	1160,39	116,39	13546,29			
16	11	1500	1270,38	124,70	1237,16	262,84	69085,69			
17	12	1610	1422,25	151,87	1395,09	214,91	46187,69			
18	13	1250	1533,15	110,90	1574,12	324,12	105055,18			
19	14	1730	1654,92	121,76	1644,05	85,95	7386,77			
20	15	1990	1803,65	148,73	1776,68	213,32	45504,62			
21	16	2030	1962,19	158,54	1952,37	77,63	6025,86			
22	17	2100	2118,11	155,92	2120,73	20,73	429,56			
23	18	1760	2209,05	90,95	2274,03	514,03	264223,21			
24	19	2300	2300,00	90,95	2300,00	0,00	0,00			
25	20	2620	2419,90	119,90	2390,95	229,05	52465,26			
26	21	2566	2543,11	123,21	2539,80	26,20	686,49			
27	22	2710	2671,84	128,73	2666,32	43,68	1907,75			
28	23	2800	2800,50	128,66	2800,58	0,58	0,33			
29	24	2850	2919,16	118,65	2929,16	79,16	6266,64			
30	25				3037,81					
31					MAD	140,91	34052,88	MSE		
32										
33					se	176,13				
34										
35										
36										
37	32		Previsão para o Mês 32		3868,383					
38										

Figura 25 Resultados após a otimização do Solver

Como você pode ver neste exemplo, o método de suavização exponencial dupla é um pouco mais complicado de se construir e deverá dar-nos resultados melhores. O que acontece se os dados mostrarem tendência como também sazonalidade? Neste caso a suavização exponencial dupla não funcionará. Precisamos usar o método da Suavização Exponencial Tripla que será discutido a seguir.

Métodos de Previsão para Séries Temporais Sujeitas a Fenômenos Sazonais e Nenhuma Tendência

É bastante comum existir padrões sazonais com valores maiores em dados instantes de tempo de que em outros em uma série temporal. Por exemplo, este fenômeno ocorre para o volume de vendas de panetones entre outros produtos típicos de festas natalinas na época do natal, assim como roupas de lã para o período de inverno, bronzeadores e bonés no período do verão, etc.

Este fenômeno viola a consideração que o processo que gera a série é por uma componente de *valor constante* ou com *tendência* e outra componente de **flutuação aleatória**, cujos métodos anteriores de previsão (*média móvel simples*, *suavização exponencial simples*, *média móvel dupla*, *suavização exponencial dupla de Holt*) não podem ser utilizados para previsão.

Uma maneira de realizar previsões com **padrões sazonais** é corrigir a série temporal do efeito da sazonalidade e, depois então, utilizar os métodos de previsão de *média móvel simples* ou *suavização exponencial simples* (para modelos de séries de valor constante⁶ e sazonalidade), como veremos nas próximas duas seções, ou ainda o método de previsão com *suavização exponencial dupla de Holt* (para modelos de séries com tendência e sazonalidade), como veremos na seção 2.2.1.4.

Considerando que o processo que gera a série temporal não tenha tendência, mas tenha sazonalidade, o modelo será dado por:

$$X_t = \pi + S_t + \varepsilon_t \quad 17$$

onde: $t = 1, 2, \dots$

X_t é uma variável aleatória observada no tempo t ;

π é o valor constante do modelo;

S_t é a componente sazonal no tempo t ;

ε_t é o erro aleatório ocorrido no tempo t (geralmente assumido ter valor esperado igual a zero e variância constante⁷).

Previsão com Correção à Priori da Sazonalidade

O procedimento pode ser resumido como:

1. Corrigir a série temporal do efeito da sazonalidade através da **divisão** (ou **subtração**) dos valores da série temporal pelos seus respectivos *fatores sazonais*.
2. Realizar a previsão através dos métodos Método de Previsão de Média Móvel Simples ou Suavização Exponencial Simples.
3. **Multiplicar** (ou **adicionar**) a previsão pelos *fatores sazonais* incorporando a sazonalidade.

Quando o efeito da sazonalidade é retirado e incorporado à série por meio de divisão e multiplicação, respectivamente, dos valores da série temporal pelos fatores sazonais o método é denominado **multiplicativo**.

Quando o efeito da sazonalidade é retirado e incorporado à série por meio de *subtração* e *adição*, respectivamente, dos valores da série temporal pelos fatores sazonais o método é denominado **aditivo**.

⁶ Sem tendência

⁷ Distribuído normalmente.

Modelo de Previsão com Sazonalidade Multiplicativa

O uso deste modelo deve ser para dados que possuam sazonalidade crescente ou decrescente, mas não possuam tendência de crescimento ou decrescimento.

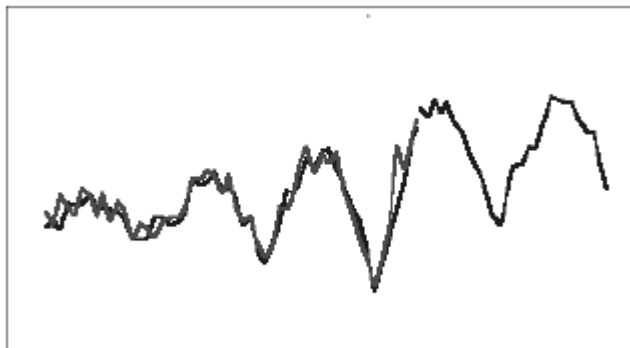


Figura 26 – Gráfico típico de dados de sazonalidade multiplicativa e sem tendência, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Este método utiliza a seguinte expressão:

$$L_t = \alpha \frac{x_t}{S_{t-s}} + (1 - \alpha) * L_{t-1} \quad 18$$

$$S_t = \gamma \frac{x_t}{L_t} + (1 - \gamma)S_{t-s} \quad 19$$

$$\hat{x}_{t+k} = L_t * S_{t+k-s} \quad 20$$

onde:

L_t é a componente de nível da série no tempo t ;

S_t é a componente de sazonalidade no tempo t ;

s é o período sazonal ou duração da sazonalidade

h é o horizonte de previsão;

$k = 1, 2, \dots, h$, isto é, o número de períodos da previsão;

\hat{x}_{t+k} é a previsão;

α com valores no intervalo $0 < \alpha < 1$, é a constante de suavização da componente de nível L_t ;

γ com valores no intervalo $0 < \gamma < 1$, é a constante de suavização da componente de sazonalidade S_t .

As constantes de suavização α e γ devem ser estimadas a partir dos dados.

Modelo de Previsão com Sazonalidade Aditiva

Este modelo pode ser usado quando ocorre sazonalidade, mas onde não se verifica a presença de tendência. Além disso, a amplitude da sazonalidade é aproximadamente constante ao longo do tempo.

O modelo pode ser usado para realizar a previsão de diversas variáveis tais como a venda de sorvetes, brinquedos, preço de commodities, etc.

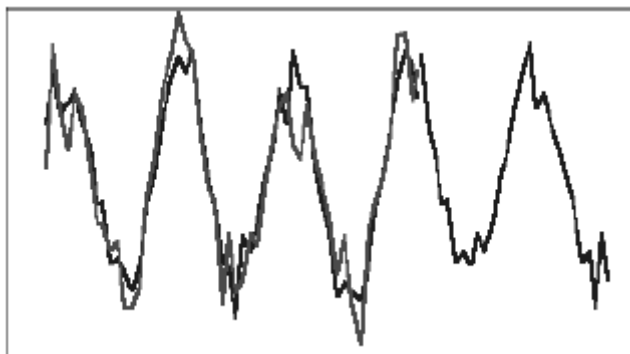


Figura 27 – Gráfico típico de dados de sazonalidade aditiva e sem tendência, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Este método utiliza a seguinte expressão:

$$L_t = \alpha * (x_t - S_{t-s}) + (1 - \alpha) * L_{t-1} \quad 21$$

$$S_t = \gamma * (x_t - L_t) + (1 - \gamma)S_{t-s} \quad 22$$

$$\hat{x}_{t+k} = L_t + S_{t+k-s} \quad 23$$

onde:

L_t é a componente de nível da série no tempo t ;

S_t é a componente de sazonalidade no tempo t ;

s é o período sazonal ou duração da sazonalidade

h é o horizonte de previsão;

$k = 1, 2, \dots, h$, isto é, o número de períodos da previsão;

\hat{x}_{t+k} é a previsão;

α com valores no intervalo $0 < \alpha < 1$, é a constante de suavização da componente de nível L_t ;

γ com valores no intervalo $0 < \beta < 1$, é a constante de suavização da componente de sazonalidade S_t .

As constantes de suavização α e γ devem ser estimadas a partir dos dados⁸.

Método de Previsão com Suavização Exponencial de Holt-Winters

Os modelos de Holt-Winters são muito utilizados quando da existência de uma série temporal que apresente, além da *tendência*, um componente de *sazonalidade*. Uma série com esse componente é caracterizada pela ocorrência de **padrões cíclicos** de variação, que se repetem em intervalos relativamente constantes de tempo. São muito observadas em indústrias do ramo alimentícias, de vestuários, cosméticos, entre outras.

Os modelos de Holt-Winters também são classificados em dois grupos: **aditivo** e **multiplicativo**. No modelo **aditivo**, a amplitude da variação sazonal é *constante* ao longo do tempo; ou seja, a diferença entre

⁸ O CB Predictor realiza isto automaticamente

o maior e menor valor de demanda dentro das estações permanece relativamente constante no tempo. No modelo **multiplicativo**, a amplitude da variação sazonal *umenta* ou *diminui* como função do tempo.

Considerando que o modelo do processo que gera a série temporal seja dado por:

$$X_t = \pi + at + S_t + \varepsilon_t \quad 24$$

Onde: $t = 1, 2, \dots$

X_t é uma variável aleatória observada no tempo t ;

a é a tendência do modelo;

π é o valor constante do modelo;

S_t é a componente sazonal no tempo t ;

ε_t é o erro aleatório ocorrido no tempo t (geralmente assumido ter valor esperado igual a zero e variância constante⁹).

Modelo Sazonal Multiplicativo de Holt-Winters

O modelo **multiplicativo** de Holt-Winters se ajusta, de maneira mais adequada, a séries com *tendência* e *sazonalidade* multiplicativa, ou seja, àquelas em que a amplitude da variação sazonal *umenta* com o acréscimo no nível médio da série temporal (KOEHLER *et al.*, 2001, p.269). Vide figura abaixo:

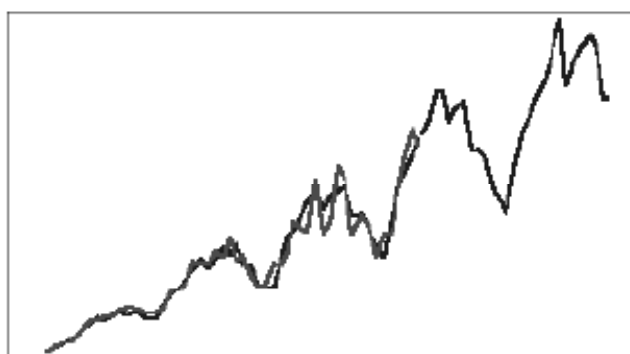


Figura 28 – Gráfico típico de dados de sazonalidade multiplicativa de Holt-Winters, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

Este método utiliza a seguinte expressão:

$$L_t = \alpha \frac{x_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad 25$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad 26$$

$$S_t = \gamma \frac{x_t}{L_t} + (1 - \gamma)S_{t-s} \quad 27$$

$$\hat{y}_{t+k} = (L_t + nT_t)S_{t-s+\text{mod}(n-1,s)+1} \quad 28$$

onde:

L_t é a componente de nível;

T_t é a componente de tendência;

⁹ Distribuído normalmente.

S_t é a componente de sazonalidade;

s é o período sazonal;

h é o horizonte de previsão;

$k = 1, 2, \dots, h$;

$\text{mod}(n,m)$ é o resto da divisão de n por m ;

\hat{y}_{t+k} é a previsão;

α com valores no intervalo $0 < \alpha < 1$, é a constante de suavização da componente de nível L_t ;

β com valores no intervalo $0 < \beta < 1$, é a constante de suavização da componente tendência T_t ;

γ com valores no intervalo $0 < \gamma < 1$, é a constante de suavização da componente de sazonalidade S_t

Na equação (25), pode-se perceber que os valores da série (x_t) são divididos pelos fatores sazonais, da mesma forma anterior em $x_{tc} = \frac{x_t}{\hat{s}_i}$ para corrigir os valores da série dos efeitos da sazonalidade, as demais parcelas da expressão são análogas as da expressão de Holt:

$$L_t = \alpha z_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad 29$$

A expressão (26) é igual à expressão (9) no método de Holt:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad 30$$

A divisão dos valores da série (y_t) pelos valores de nível (L_t) na expressão (21) pode ser entendida como a medida de fator sazonal “instantânea”.

Em (28) a sazonalidade é incorporada à série através da multiplicação da soma dos valores previstos para as componentes de Nível (L_t) e Tendência (T_t) pela componente sazonal S_{t-s+k} .

O método **multiplicativo** de Winters, como os demais modelos descritos anteriormente, funciona através da aplicação **recursiva** de suas equações aos dados da série. Dessa forma, tal aplicação deve iniciar em algum período no passado, onde os valores de L_t , T_t e S_t devem ser estimados (MAKRIDAKIS et al., 1998, p.168). Uma maneira simples de se fazer essa estimativa é através da inicialização do *nível* e da *tendência* no mesmo período m :

- O **nível** é determinado através da média de primeira estação:

$$L_s = \frac{1}{s} (x_1 + x_2 + \dots + x_s) \quad 31$$

- Para se inicializar a **tendência**, é recomendado o uso de duas estações completas, ou seja, 2s períodos:

$$T_s = \frac{1}{s} \left(\frac{x_{s+1} - x_1}{s} + \frac{x_{s+2} - x_2}{s} + \dots + \frac{x_{s+s} - x_s}{s} \right) \quad 32$$

- Por último, os índices sazonais iniciais podem ser determinados através da razão entre as primeiras observações com a média do primeiro ano:

$$S_1 = \frac{x_1}{L_s}, S_2 = \frac{x_2}{L_s}, \dots, S_s = \frac{x_s}{L_s} \quad 33$$

Nas expressões acima, \hat{y}_{t+k} é a previsão para o período $t+k$, α , β e γ são constantes de suavização, cujos valores encontram-se entre 0 e 1, e y_t é a mais recente observação. Nelas, temos:

- L_t representa uma estimativa do nível da série no tempo t ,
- T_t , uma estimativa da declividade da série no mesmo período t e
- S_t , o componente de sazonalidade também no período t .

A sazonalidade, ou seja, o número de subperíodos por ano, é representado por s . A escolha dos valores para as constantes de suavização α , β e γ é condicionada a algum critério que, na maioria das vezes, consiste no mesmo citado anteriormente: a minimização pelo uso de um algoritmo de otimização não linear, do erro quadrático médio (EQM) atribuído ao desempenho do modelo usando a ferramenta Solver do MS-Excel.

Vamos começar a construir o modelo. Vá à planilha *Solução EXP TRIPLA* na pasta *Séries Temporais*.

	A	B	C	D	E	F	G	H	I	J	K	L
2			y_t	E_t	T_t	s_t	\hat{y}_t	$ e_t $	e_t^2			
3		Mês	Vendas de Varas	Média Suavizada	Tendência	Fatores Sazonais	Previsto	erro absoluto	erro quadrático			
4	jan	1	7				0,21					
5	fev	2	5				0,15					
6	mar	3	15				0,45					
7	abr	4	25				0,74					
8	mai	5	42				1,25					
9	jun	6	48				1,43					
10	jul	7	70				2,08					
11	ago	8	75				2,23					
12	set	9	40				1,19					
13	out	10	30				0,89					
14	nov	11	25				0,74					
15	dez	12	22	33,67	0,00	0,65						
16	jan	13	10	40,88	3,61	0,23	7,00	3,00	9,00		Alfa	0,50
17	fev	14	7	45,81	4,27	0,15	6,61	0,39	0,15		Beta	0,50
18	mar	15	20	47,48	2,97	0,43	22,31	2,31	5,35		Gama	0,50
19	abr	16	32	46,77	1,13	0,71	37,47	5,47	29,88			
20	mai	17	58	47,20	0,78	1,24	59,76	1,76	3,11			
21	jun	18	60	45,03	-0,70	1,38	68,40	8,40	70,58			
22	jul	19	90	43,81	-0,96	2,07	92,18	2,18	4,75			
23	ago	20	95	42,75	-1,01	2,23	95,46	0,46	0,21			
24	set	21	60	46,12	1,18	1,24	49,59	10,41	108,36			
25	out	22	40	46,09	0,58	0,88	42,15	2,15	4,62			
26	nov	23	37	48,25	1,37	0,75	34,66	2,34	5,49			
27	dez	24	30	47,76	0,44	0,64	32,42	2,42	5,87			
28	jan	25	16	59,46	6,07	0,25	10,91	5,09	25,94			
29	fev	26	20	99,14	22,87	0,18	9,87	10,13	102,58			
30	mar	27	50	118,69	21,21	0,43	52,88	2,88	8,27			
31	abr	28	80	126,03	14,27	0,67	99,80	19,80	392,21			
32	mai	29	150	130,72	9,49	1,19	173,72	23,72	562,61			
33	jun	30	152	125,21	1,99	1,30	193,36	41,36	1710,70			
34	jul	31	235	120,45	-1,39	2,01	262,89	27,89	777,99			
35	ago	32	250	115,71	-3,06	2,19	264,92	14,92	222,66			
36	set	33	175	126,63	3,93	1,31	140,20	34,80	1211,29			
37	out	34	110	127,82	2,56	0,87	114,82	4,82	23,22			
38	nov	35	100	131,44	3,09	0,76	98,40	1,60	2,57			
39	dez	36	80	129,69	0,67	0,63	86,20	6,20	38,49			
40	jan	37					32,29					
41								9,77	221,91			
42								MAD	MSE			
43												
44								12,21				
45								se				
46												
47	ago	44					296,11					
48												
49												

Figura 29

Aqui temos as vendas mensais de uma empresa de comércio de varas de pescar. O pico das vendas durante a estação de pesca no outono e verão. Portanto sabemos que é sazonal anualmente onde $c = 12$. O primeiro passo é transformar ou desazonalizar os dados das vendas do primeiro ano no intervalo C4:C15. Assim entramos com a fórmula $=C4/((1/12)*SOMA(\$C\$4:\$C\$15))$ na célula F4 e preencha até a F15. Isto é realmente como

$$s_t = \frac{y_t}{\left[\left(\frac{1}{c}\right)(y_1 + y_2 + \dots + y_c)\right]}$$

é entrado no Excel.

Em segundo lugar entramos com a fórmula: $E_t = \alpha (y_t/S_{t-c}) + (1 - \alpha) (E_{t-1} + T_{t-1})$ na célula D16 que é $=\$J\$16*C16/F4+(1-\$J\$16)*(D15+E15)$ e preencha até a D39.

Depois disto entramos com esta fórmula: $T_t = \beta (E_t - L_{t-1}) + (1 - \beta) T_{t-1}$ na célula E16 que é $=\$J\$18*(D16-D15)+(1-\$J\$18)*E15$ e preencha até E39.

O próximo passo é entrar com a fórmula: $S_t = \gamma (y_t/L_t) + (1 - \gamma) S_{t-c}$ na célula F16 que é $=\$J\$20*C16/D16+(1-\$J\$20)*F4$ e preencha até a F39

Por último, entre com a fórmula: $\hat{y}_{t+1} = (E_t + T_t)S_{t+1-c}$ na célula G16 que é $=(D15+E15)*F4$ e preencha até G39.

O erro absoluto e o erro quadrado são entrados na coluna H e I, respectivamente. A célula H41 mostra o MAD e a célula I41 é o MSE. Usaremos o Solver do Excel para minimizar o MSE e encontrar o melhor valor ou otimizar α , β e γ . Inicialmente usamos $\alpha = 0,5$, $\beta = 0,5$ e $\gamma = 0,5$ indicados nas células J16, J18 e J20, respectivamente (ver Figura 29).

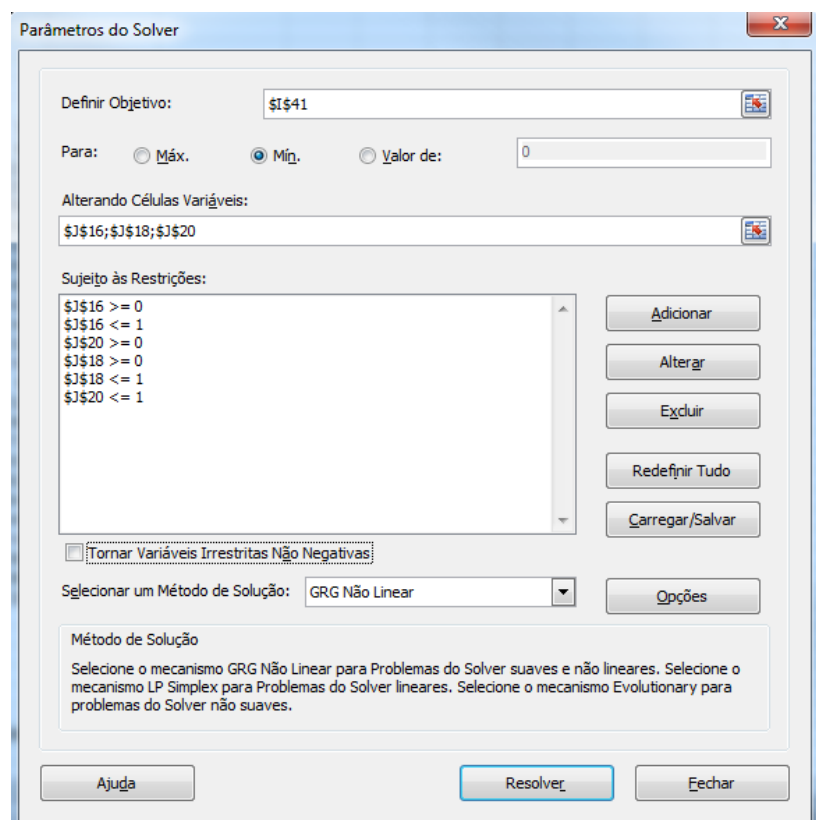


Figura 30

Entramos com todos os parâmetros nesta caixa de diálogo. Ela se parecerá como a Figura 30 acima após todos os parâmetros terem sido entrados.

Definir objetivo: MSE(I41); Para: Min

Alterando Células Variáveis: os pesos em J16 para α , J18 para β e J20 para γ

Sujeito às Restrições: $0 \leq J16 \leq 1$, $0 \leq J18 \leq 1$ e $0 \leq J20 \leq 1$

Clique no botão Resolver. O Solver iniciará a otimização.

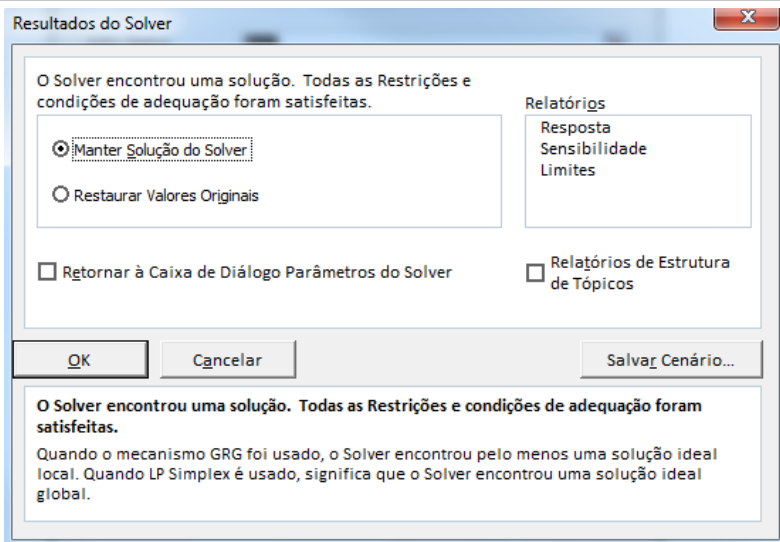


Figura 31

Mantenha a solução encontrada pelo Solver.

O MSE na célula I41 é agora 49,17, comparado com 221,91, inicialmente. O $\alpha = 0,46$ na célula J16, $\beta = 0,05$ na célula J18 e $\gamma = 1$ na célula J20 (ver Figura 32 abaixo).

	A	B	C	D	E	F	G	H	I	J	K
2			yi	Ei	Ti	si	yhati	leij	ei^2		
3		Mês	Vendas de Varas	Média Suavizada	Tendência	Fatores Sazonais	Previsto	erro absoluto	erro quadrático		
4	jan	1	7			0,21					
5	fev	2	5			0,15					
6	mar	3	15			0,45					
7	abr	4	25			0,74					
8	mai	5	42			1,25					
9	jun	6	48			1,43					
10	jul	7	70			2,08					
11	ago	8	75			2,23					
12	set	9	40			1,19					
13	out	10	30			0,89					
14	nov	11	25			0,74					
15	dez	12	22	33,67	0,00	0,65					
16	jan	13	10	40,27	0,31	0,25	7,00	3,00	9,00	0,46	Alfa
17	fev	14	7	43,58	0,45	0,16	6,03	0,97	0,95	0,95	Beta
18	mar	15	20	44,42	0,46	0,45	19,61	0,39	0,15	0,05	0,05
19	abr	16	32	44,06	0,43	0,73	33,33	1,33	1,77	1,77	Gama
20	mai	17	58	45,41	0,47	1,28	55,50	2,50	6,23	6,23	1,00
21	jun	18	60	44,14	0,39	1,36	65,41	5,41	29,24	29,24	
22	jul	19	90	43,96	0,36	2,05	92,59	2,59	6,68	6,68	
23	ago	20	95	43,55	0,33	2,18	98,74	3,74	13,97	13,97	
24	set	21	60	46,91	0,47	1,28	52,14	7,86	61,85	61,85	
25	out	22	40	46,24	0,41	0,87	42,22	2,22	4,91	4,91	
26	nov	23	37	48,10	0,48	0,77	34,64	2,36	5,56	5,56	
27	dez	24	30	47,36	0,42	0,63	31,75	1,75	3,06	3,06	
28	jan	25	16	55,40	0,78	0,29	11,87	4,13	17,08	17,08	
29	fev	26	20	87,44	2,23	0,23	9,02	10,98	120,46	120,46	
30	mar	27	50	99,45	2,69	0,50	40,38	9,62	92,61	92,61	
31	abr	28	80	105,81	2,86	0,76	74,18	5,82	33,93	33,93	
32	mai	29	150	112,68	3,04	1,33	138,81	11,19	125,32	125,32	
33	jun	30	152	113,94	2,96	1,33	157,30	5,30	28,07	28,07	
34	jul	31	235	115,93	2,92	2,03	239,32	4,32	18,69	18,69	
35	ago	32	250	116,91	2,83	2,14	259,22	9,22	85,06	85,06	
36	set	33	175	127,55	3,19	1,37	153,15	21,85	477,47	477,47	
37	out	34	110	129,10	3,11	0,85	113,10	3,10	9,62	9,62	
38	nov	35	100	131,21	3,07	0,76	101,69	1,69	2,86	2,86	
39	dez	36	80	130,62	2,90	0,61	85,05	5,05	25,52	25,52	
40	jan	37					38,56				
41								5,27	49,17		
42								MAD	MSE		
43											
44								6,58			
45								se			
46											
47	ago	44					328,86				
48											
49											

Figura 32

Fizemos um grande melhora após a otimização com o Solver do Excel.

Se no mês 36 vendermos 80 varas de pescar, como você faria para prever a venda no mês 37? Para prever simplesmente preencha a fórmula da célula G39 na célula G40 ou entre com a fórmula $= (D39 + E39) * F28$. Aqui o resultado é 38,56 varas ou arredondando 39.

Se você quiser prever k períodos adiante, use a seguinte lógica:

$$\hat{y}_{t+k} = (E_t + kT_t)s_{t+k-c}$$

Neste exemplo queremos prever as vendas no mês 44, i.e., 8 meses adiante, então entramos com $= (D39 + 8 * E39) * F35$ como você pode ver na célula G47. O resultado é 328,86, ou arredondando 329 varas de pescar (ver Figura 1.21 acima).

$$E_1 = D39 \quad T_1 = E39 \quad s_{t+k-c} = F35 \quad k = 8 \quad c = 12$$

Como você pode ver neste exemplo, o método da suavização exponencial tripla é um pouco mais complicado de se construir e nos dará resultados muito bons.

Modelo Sazonal Aditivo de Holt-Winters

Para séries que possuem *tendência* e *sazonalidade aditiva*, o modelo que apresenta maior capacidade de explicação é o **aditivo de Winters**. Ou seja, ele é utilizado nas séries onde o *efeito sazonal não* é função do nível médio corrente da série temporal e pode ser adicionado ou subtraído de uma previsão que dependa apenas de *nível* e *tendência* (KOHLENER *et al.*, 2001, p. 269). Veja figura abaixo:

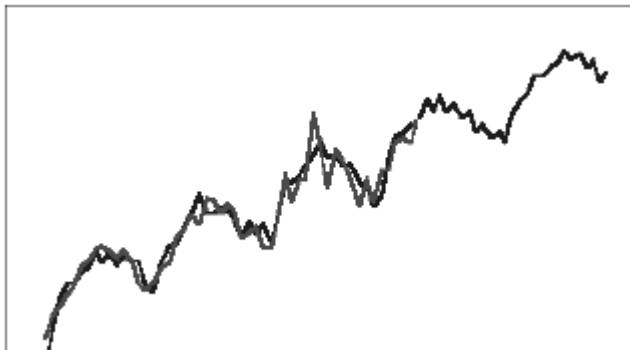


Figura 33 – Gráfico típico de dados de sazonalidade aditiva de Holt-Winters com tendência, mostrando o ajuste e a linha de previsão

Fonte: CB Predictor – *User Manual* – Selecionando métodos de previsão de séries temporais. P. 3-10

O algoritmo de previsão do modelo sazonal **aditivo de Holt-Winters** é baseado nas seguintes expressões:

$$L_t = \alpha(x_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad 34$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad 35$$

$$S_t = \gamma(x_t - L_t) + (1 - \gamma)S_{t-s} \quad 36$$

$$\hat{x}_{t+k} = L_t + kT_t + S_{t-s+k} \quad 37$$

Estas equações são parecidas com aquelas a que se refere ao modelo multiplicativo de Holt-Winters (Eqs. 25 a 28). A diferença nos dois modelos é o fato das outras equações, agora, apresentarem os índices de sazonalidade *somados* e *subtraídos*, ao invés de multiplicados e divididos (MAKRIDAKIS *et al.*, 1998, p.169).

As inicializações de L_s e T_s são idênticas às do modelo multiplicativo. Os valores iniciais para os índices sazonais são determinados através das seguintes expressões:

$$S_1 = x_1 - L_s, S_2 = x_2 - L_s, \dots, S_s = x_s - L_s$$

Conclusão

A média móvel é comumente usada com dados de séries temporais para suavizar flutuações de curto prazo e destacar tendências ou ciclos de longo prazo. O corte entre curto prazo e longo prazo depende da aplicação e os parâmetros da média móvel serão definidos apropriadamente. Por exemplo, é frequentemente usado em análise técnica de dados financeiros, como preços de ações, retornos ou volumes transacionados. É usado também em economia para examinar o produto interno bruto, empregos e outras séries temporais macroeconômicas.

A suavização exponencial tem provado através dos anos como muito útil em muitas situações de previsão. Ela foi sugerida primeiramente por C.C.Holt em 1957 e tinha a intenção de ser usada para séries temporais não sazonais não mostrando tendência. Mais tarde ele ofereceu um procedimento (1958) que manipula tendências. Winters (1965) generalizou o método para incluir sazonalidade, daí o nome “Método de Holt-Winters” ou Suavização Exponencial Tripla.

Todos estes métodos de previsão são muito básicos mas muito úteis. Os métodos de previsão de séries temporais podem ser mais avançados do que estes considerados nos nossos exemplos acima. Eles são baseados nos modelos *Auto Regressive Integrated Moving Average* (ARIMA) (também conhecido como técnica de Box-Jenkins). Essencialmente estes assumem que as séries temporais foram geradas por um processo probabilístico com valores futuros relacionados aos valores passados, como também aos erros de previsão passados. Para aplicar os modelos ARIMA a série temporal precisa ser estacionária. Uma série temporal estacionária é aquela uma cujas propriedades estatísticas tais como média, variância e autocorrelação são constantes no decorrer do tempo. Aprenderemos como modelar com ARIMA como um método de previsão noutro texto: Técnicas de Previsão de Box-Jenkins no Excel

Apêndice A – Critérios para avaliar o desempenho de modelos em relação à série temporal

Para saber se o modelo de séries temporais selecionado para realizar previsões é significativo estatisticamente devemos quantificar o ajuste do modelo escolhido aos dados históricos existentes. Para isso, há diversas estatísticas:

Medidas de erros no processo de previsão

MAD (Média dos Desvios Absolutos)

Consiste na média da diferença (em módulo) entre os valores reais e preditos:

$$MAD = \frac{1}{n} \sum_{t=1}^n |e_t| \quad 39$$

Ela mede o erro absoluto e originalmente tornou-se muito popular (anteriormente às calculadoras) porque não exige cálculos de quadrados e raiz quadrada. Por enquanto é ainda muito confiável e amplamente utilizada, é muito acurada para dados distribuídos normalmente.

MAPE (Média Aritmética dos Desvios Absolutos em Porcentuais)

Avalia a magnitude do erro com relação à série histórica:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\varepsilon_t}{z_t} \times 100 \right| \quad 40$$

É a mais popular das formas apresentadas para medir a acurácia da previsão (Kahn, 1998). Porém, o uso desta fórmula torna-se impossível quando a série temporal contém valores iguais à zero.

Por se tratar de erro relativo, ele não depende da escala, e com isso, permite comparar a acurácia da previsão entre séries de dados temporais de proporções diferentes.

EQD (Média Aritmética dos Quadrados dos Desvios) ou MSE

Destaca os grandes erros, comparados aos erros de menor magnitude:

$$EQM = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad 41$$

RMSE (Raiz Quadrada da Média Aritmética dos Quadrados dos Desvios)

É a raiz quadrada dos EQM's:

$$RMSE = \sqrt{EQD} \quad 42$$

Esta medida também tende a exagerar erros grandes, que podem ajudar a eliminar métodos com grandes erros.

R²

Existe outro parâmetro, além dessas medidas de erro, de grande importância na análise do ajuste do modelo à série temporal. Trata-se do R², uma medida percentual de explicação do modelo. Ou seja, ele relata a fração da variabilidade da série que o método utilizado consegue explicar. Um valor R² próximo de 0 indica um modelo de ajuste pobre, enquanto um valor próximo a 1 indica um bom ajuste. Essa estatística é determinada da seguinte maneira:

$$R^2 = 1 - \frac{\sum_{t=1}^n (x_t - \hat{x}_t)^2}{\sum_{t=1}^n (x_t - \bar{x}_t)^2} \quad 43$$

Medidas estatísticas de qualidade de previsão

U de Thiel

É uma estatística que permite estimar a probabilidade de que o modelo selecionado é melhor do que simplesmente uma previsão com base em uma simples média aritmética:

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{x}_{t+1} - x_{t+1}}{x_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{x_{t+1} - x_t}{x_t} \right)^2}}$$

onde:

N é o tamanho da amostra

$x(t)$ são os valores históricos

\hat{x}_t representa os valores previstos

Interpretação do U :

$U < 1$: a previsão pelo método escolhido é melhor do que simples adivinhação (“chute”);

$U = 1$: a previsão pelo método escolhido é tão boa ou ruim quanto uma simples adivinhação;

$U > 1$: a previsão pelo método escolhido é pior que uma adivinhação.

Esta medida também exagera os erros e ajuda eliminar métodos com grandes erros.

Teste de Durbin-Watson

É uma estatística que permite analisar a possível presença de auto-correlação da variável para uma defasagem de uma unidade ($lag1$).

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

onde:

n é o tamanho da amostra;

$e(t)$ se refere aos resíduos entre os valores previstos e pelo método selecionado e os valores reais.

Interpretação de D-W:

$DW < 1$: os valores da série histórica são positivamente correlacionados;

$DW = 2$: não há auto-correlação entre os valores da série histórica;

$DW > 3$: os valores da série são negativamente correlacionados.

Os demais valores devem ser interpretados de forma proporcional.

Estatística de Ljung-Box

É uma grandeza que permite analisar se um conjunto de auto-correlações com valores diferentes de zero é significativamente diferente de um conjunto de auto-correlações com valores iguais a zero

$$Q' = n(n+2) \sum_{k=1}^{h-1} \frac{r_k^2}{(n-k)} \quad 46$$

onde:

n é o número de pares de pontos;

h é o tamanho do conjunto de auto-correlações

k é o tamanho das defasagens e r_k são os valores das auto-correlações de ordem k .

Intervalo de Confiança na previsão

O intervalo de confiança da previsão pode ser interpretado como uma medida de precisão do modelo de previsão. O erro padrão de previsão é estimado por meio de:

$$\text{Erro - padrão} = \sqrt{\frac{\sum_{t=1}^{n-m} e_t^2(m)}{n-m}} \quad 47$$

É importante notar que se a série de dados históricos for N, o número máximo de intervalos a serem previstos será igual a N-1.

BIBLIOGRAFIA

- ARCHER, B. Forecasting Demand: Quantitative and Intuitive Techniques. **International Journal of Tourism Management**. USA. v.1, n.1, p. 5-12, 1980.
- BOX,G.; JENKINS, G. **Time series analysis: Forecasting and control**. San Francisco: Holden-Day, 1970.
- CHAMBERS, J. C.; MULLICK, S. K.; SMITH, D. D. How to Choose the Right Forecasting Technique. **Harvard Business Review**. USA. v. 49, July-August, p. 45-57, 1971.
- CHATFIELD, C. **The Analysis of Time Series**. 6 ed. Washington: Chapman & Hall/CRC, 2004. 322p.
- DAVIS, M.; AQUILANO, N.; CHASE, R. **Fundamentos da Administração da Produção**. 3 ed. Porto Alegre: Bookman, 2001. 472p.
- EHLERS, R.S. Análise de Séries Temporais. Disponível em: < <http://www2.icmc.usp.br/~ehlers/>>. Acesso em 22 de dez 2012
- FERREIRA, A.B.H. Novo Dicionário da Língua Portuguesa. Rio de Janeiro: Editora Nova Fronteira, 1975. 1516p.
- FURTADO, M.R. **Aplicação de um modelo de previsão da demanda total nos credenciados Belgo pronto**. 2006. 32 p. Monografia (Conclusão de Graduação em Engenharia de Produção) - Curso de Engenharia de Produção, Universidade Federal de Juiz de Fora, Juiz de Fora, 2006.
- HARRISON, P.J.; STEVENS, C. F. Bayesian Forecasting. **Journal of the Royal Statistical Society**. UK. series B, v. 38, nº 3, p.81-135, 1976
- KOEHLER, A. B.; SNYDER, R. D.; ORD, J. K. Forecasting Models and Prediction Intervals for the Multiplicative Holt-Winters Method. **International Journal of Forecasting**. USA. v.17, n.2, p.269-286, apr./jun. 2001.
- MAKRIDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R. **Forecasting: Methods and Applications**. 3. ed., New York: John Wiley & Sons, 1998.
- MONTGOMERY, D.; JOHNSON, L.; GARDINER, J. **Forecasting and Time Series Analysis**. New York: McGraw-Hill, 1990.
- MORETTIN, P.A.; TOLOI, C.M.C. **Análise de Séries Temporais**. São Paulo: Editora Edgard Blucher Ltda., 2004.
- MUN, J. **Modeling Risk: Applying Monte Carlo Simulation, Real Options Analysis, Stochastic Forecasting and Portfolio Optimization**. 2ª ed., New York:John Wiley & Sons, 2010.
- ORACLE CRYSTAL BALL – **CBPredictor User Manual**, Release 11.1.1.1.00, 2008
- PINDYCK, R.; RUBENFELD, D.L. **Microeconomia**. 5ª ed., São Paulo: Prentice Hall, 2002.
- TSAY, R.S. **Analysis of Financial Times Series: Financial Econometrics**. New York: John Wiley & Sons, 2002