Data Analytics

Time Series - Deep Dive : Decomposition Methods , Exponential Smoothing

Dr. Rita Chakravarti
Institute of Systems Science
National University of Singapore
Email: rita@nus.edu.sg

Quantitative Forecasting Techniques

Causal

- Regression Analysis

Univariate

- Time Series Regression (t is the independent variable)
- Classical Decomposition Methods
- Exponential Smoothing (2 & 3 parameters)
- Box-Jenkins Methods

Classical Decomposition of Time Series

- **Trend** – does not necessarily imply a monotonically increasing or decreasing series but simply a lack of constant mean, though in practice, a linear or quadratic function is often to predict the trend;
- **Cycle** – refers to patterns or waves in the data that are repeated after approximately equal intervals with approximately equal intensity. For example, some economists believe that “business cycles” repeat themselves every 4 or 5 years;
- **Seasonal** – refers to a cycle of one year duration;
- **Random (irregular)** – refers to the (unpredictable) variation not covered by the above

Decomposition Methods

Decomposition Methods are used to forecast time series that exhibit trend and seasonal effects

These models have no theoretical basis – they are strictly an intuitive approach

They are useful when the parameters describing a time series is not changing over time

The basic idea behind these models is to decompose the time series into several factors : 1) Trend 2) Seasonal 3) Cyclical and 4) Irregular (error)

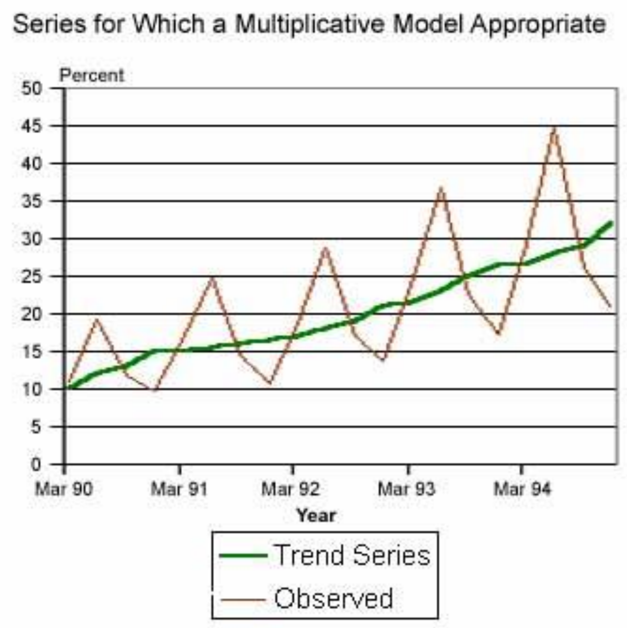
Two Types of Decomposition Methods

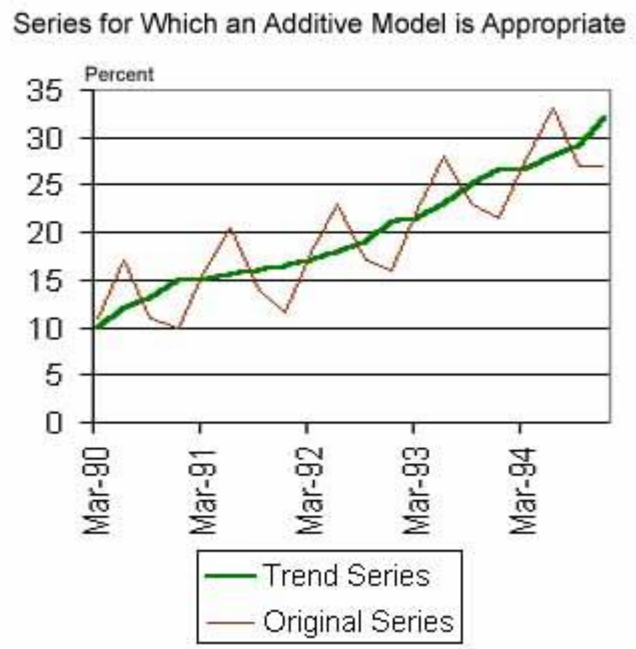
Multiplicative Models

$$Y_t = TR_t \times SN_t \times CL_t \times IR_t$$

Additive Models

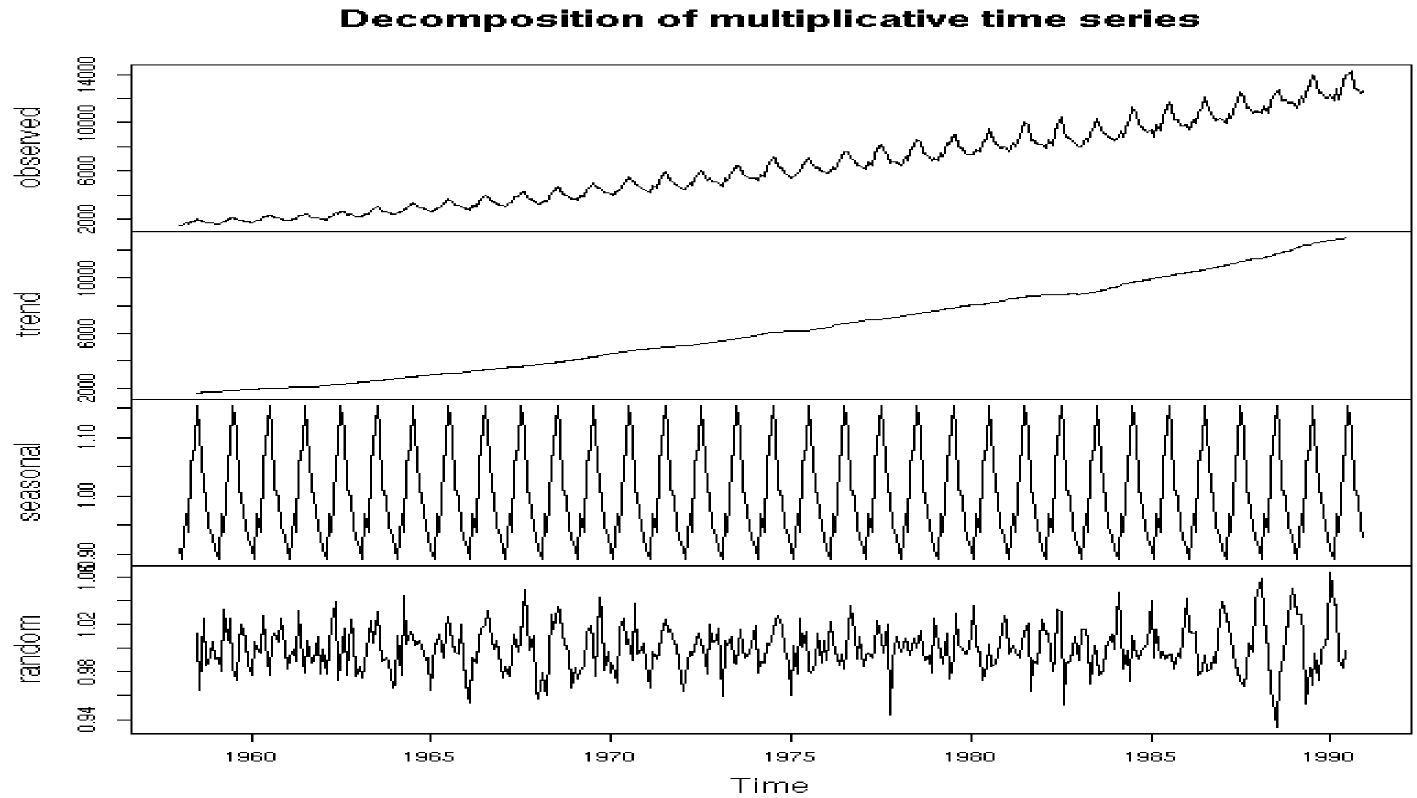
$$Y_t = TR_t + SN_t + CL_t + IR_t$$





Multiplicative Decomposition

Consider a Time Series that exhibits increasing or decreasing Seasonal Variation



Multiplicative Decomposition Method

Multiplicative Models

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

Where.

y_t = the observed value of the time series at time t

TR_t = the trend component (or factor) at time t

SN_t = the seasonal component (or factor) at time t

CL_t = the cyclical component (or factor) at time t

IR_t = the irregular component (or factor) at time t

Multiplicative Decomposition Method

This decomposition method employs multiplicative seasonal factor

i.e. seasonal factor is multiplied by trend (instead of added to the trend like dummy variable regression)

Example Sales (in case the model is multiplicative) :

Let the trend equation be : $TR_t = 500 + 50t$,   t =0, Q_4 2002

So in 2003 the forecasts using trend equation only is :

$TR_1 = 550$ Q_1 2003 $TR_2 = 600$ Q_2 2003

$TR_3 = 650$ Q_3 2003 $TR_4 = 700$ Q_4 2003

Suppose it is known that sales are lowest in the Q_1 highest in the $Q_{2,}$ moderately high in the Q_3 and moderately low in Q_4

As we know the sales are seasonal and past record shows that

$SN_{Q1} = 0.4$ $SN_{Q2} = 1.6$ $SN_{Q3} = 1.2$ $SN_{Q4} = 0.8$

Multiplicative Decomposition Method

So assuming multiplicative model, if we consider both trend and seasonal effects (trend * seasonal)

So in 2003 the forecasts using trend and seasonality would be :

$TR_1 X SN_{Q1} = 220$ Q_1 2003 $TR_2 X SN_{Q2} = 960$ Q_2 2003

$TR_3 X SN_{Q3} = 780$ Q_3 2003 $TR_4 X SN_{Q4} = 560$ Q_4 2003

If the multiplicative seasonal factors remain constant over time what would be the forecasts in 4 quarters in 2004?

Multiplication of the trend by seasonal factors implies that the size of the seasonal swing is proportional to the trend. So if the trend increases so does the swing.

However the assumption is that the seasonal factors are constant over time. If not other techniques need to be used.

Multiplicative Decomposition Method

The seasonal factor SN_t models cyclical pattern in a time series that are completed within a calendar year. If a time series displays a cycle that has a longer duration, a cyclical factor can be defined. Suppose in our sales data all 4 quarters of 2003 are included in the boom period of a business cycle. Also suppose $CL_1 = 1.08$, $CL_2 = 1.09$, $CL_3 = 1.09$, $CL_4 = 1.10$. If we consider trend, seasonal & cyclical factors, sales in the four quarters of 2003 would be :

$TR_1 X\, SN_{Q1}\, X\, CL_1 = 238$ Q_1 2003 $TR_2 X\, SN_{Q2}\, X\, CL_2 = 1046$ Q_2 2003

$TR_3 X\, SN_{Q3}\, X\, CL_3 = 850$ Q_3 2003 $TR_4 X\, SN_{Q4}\, X\, CL_4 = 616$ Q_4 2003

If the multiplicative cyclical factors in 2004 are 1.09, 1.05, 1.01, 0.99 in four quarters what would be the forecasts in 4 quarters in 2004?

Multiplicative Decomposition Method

Work shop

- A cola company owns and operates ten drive in soft drink stores. They were selling "diet cola" that was introduced in the market just three years ago and has been gaining popularity. They order the supply of the "diet cola" from a regional distributor. The company has an inventory policy that attempts to meet practically all of the demand for "diet cola", while at the same time ensuring that the company does not tie up its money needlessly by ordering much more than the actual demand. In order to implement its inventory policy the company needs to forecast monthly "diet cola" sales(in hundreds of cases).
- At the end of each month the cola company desires point forecasts and prediction interval forecasts of "diet cola" in future months
- The company has recorded monthly "diet cola" sales for previous 3 years in the data set "Cola_data.xls"

Multiplicative Decomposition Method

Step by step approach

1) Calculate **12** month moving average (MA_t) - to eliminate **seasonal** & irregular fluctuation --- Question : If it was quarterly data what would you do???

2) Next calculate Centered Moving Average CMA_t :Why??

Clue : we are calculating moving average of even order. If it was of odd order, this step was not necessary

Since the model is $y_t = TR_t \times SN_t \times CL_t \times IR_t$, this implies that

$$SN_t \times IR_t = y_t / TR_t \times CL_t = y_t / CMA_t$$

Because CMA_t is considered to be an estimate of $TR_t \times CL_t$, as averaging process is assumed to have removed seasonal and short term irregular fluctuations

3) Calculate $SN_t \times IR_t$

Multiplicative Decomposition Method

Step by step approach

4) Next step calculate average SN_t for each season (Jan, Feb, ….)

5) Then find normalization factor for SN_t . As L=12 (number of seasons) the normalization factor is $12/\sum_1^L SN_t$

6) By multiplying average SN_t by normalizing factor one would get estimate of SN_t.

7) Next step is to calculate the de-seasonalized observation d_t at time t:

$$d_t = y_t / \text{ estimated } SN_t$$

Deseasonalized observations are computed in order to better estimate the trend components. By dividing y_t by seasonality we can have better understanding of trend

8) Next step is to estimate TR_t . Look at plot of d_t , if it is linear please fit

$$TR_t = \beta_0 + \beta_1 t$$

This completes the estimation of SN_t (step 6) and TR_t

Multiplicative Decomposition Method

Step by step approach

9) So from the equation $y_t = TR_t \times SN_t \times CL_t \times IR_t$

it implies that $CL_t \times IR_t = y_t / (TR_t \times SN_t)$ where one can use the estimate of TR_t and SN_t

10) It is observed empirically that when considering monthly or quarterly data we can average out ir_t by taking a three period moving average

11) Finally we calculate the estimate of IR_t by using the formula,

$$cl_t \ X \ ir_t \ / \ cl_t$$

Note : Traditionally the estimates tr_t, sn_t, cl_t and ir_t are obtained by using multiplicative decomposition method and used to describe the time series.

However we can also use these estimates to forecast the future values of time series. If there is no pattern in irregular component, we predict IR_t to be 1.

Multiplicative Decomposition Method

Forecasting using the method: In case $IR_t = 1$,

A) The point forecast of $y_t = tr_t \times sn_t \times cl_t$, if a well defined cycle exists and can be predicted

B) The point forecast of $y_t = tr_t \times sn_t$, if a well defined cycle does not exist or if CL_t can not be predicted

For our Cola example where

$$tr_t = b_0 + b_1 t$$

We can find the forecast value of TR_t by taking $t = 37, 38, \ldots\ldots$ Multiply each TR_t by corresponding monthly SN_t to obtain the point estimate

Although there is no theoretically correct prediction interval for y_t Bowerman , O'Connell & Koehler have found that a fairly accurate (approx.) $100(1-\alpha)\%$ prediction interval for y_t is

$[y_t \pm B_t[100(1-\alpha)]]$ where $B_t[100(1-\alpha)]$ is a error bound in a $100(1-\alpha)\%$ prediction interval $[tr_t \pm B_t[100(1-\alpha)]]$ for the deseasonalized observation

$$d_t = TR_t + \varepsilon_t = \beta_0 + \beta_1 t + \varepsilon_t$$

Multiplicative Decomposition - Sales

$$Y_t = TC_t \times SN_t \times IR_t$$

Examples:

Quarterly sales from 2010 to 2013 – Assumption no cyclical trend

Period (t)	Year	Quarter	Sales
1	1	1	72
2		2	110
3		3	117
4		4	172
5	2	1	76
6		2	112
7		3	130
8		4	194
9	3	1	78
10		2	119
11		3	128
12		4	201
13	4	1	81
14		2	134
15		3	141
16		4	216

Estimation of seasonal component (SN_t)

$$Y_t = TC_t \times SN_t \times IR_t$$

$$\therefore \qquad \hat{S}N_t = \frac{Y_t}{TC_t \times IR_t}$$

Follow similar steps 1, ….9 articulated on slides 12-14 , adjusting for quarterly (and not monthly) data and ignoring CL_t

Calculation of MA & CMA (steps 1-2)

Period (t)	Year	Quarter	Sales	MA (4)	CMA
1	1	1	72		
2		2	110		
3		3	117	T=2.5 117.75	T=3 118.25
4		4	172	T=3.5 118.75	119
5	2	1	76	119.25	120.875
6		2	112	122.5	125.25
7		3	130	128	128.25
8		4	194	128.5	129.375
9	3	1	78	130.25	130
10		2	119	129.75	130.625
11		3	128	131.5	131.875
12		4	201	132.25	134.125
13	4	1	81	136	137.625
14		2	134	139.25	141.125
15		3	141	143	
16		4	216		

Calculating SN_t (steps 3-6)

Quarter	Sales	MA(4)	TR	SNXIR(t)	Avg(SN(t))	Exp(SN(t))
1	72				0.606	0.60661
2	110				0.918	0.91892
3	117	117.75	118.25	0.989	0.991	0.99199
4	172	118.75	119	1.445	1.481	1.48248
1	76	119.25	120.875	0.629	0.606	0.60661
2	112	122.5	125.25	0.894	0.918	0.91892
3	130	128	128.25	1.014	0.991	0.99199
4	194	128.5	129.375	1.5	1.481	1.48248
1	78	130.25	130	0.6	0.606	0.60661
2	119	129.75	130.625	0.911	0.918	0.91892
3	128	131.5	131.875	0.971	0.991	0.99199
4	201	132.25	134.125	1.499	1.481	1.48248
1	81	136	137.625	0.589	0.606	0.60661
2	134	139.25	141.125	0.95	0.918	0.91892
3	141	143			0.991	0.99199
4	216				1.481	1.48248
					3.996	
					1.001001001	

S1. TR(3) = 117/118.25=0.9894

S2. Average SN(t)
= average(yr1, yr2, yr3)

S3. Normalizing factor
= 4 /sum(average(SN(t))

S4. Expected SN(t)
= S2 * S3

Computation of d_t (step 7)

Period(t)	Year	Quarter	Sales	Exp(SN(t))	d=deseasonalize Y	
1	1	1	72	0.60661	118.69	
2		2	110	0.91892	119.71	
3		3	117	0.99199	117.94	
4		4	172	1.48248	116.02	
5	2	1	76	0.60661	125.29	
6		2	112	0.91892	121.88	
7		3	130	0.99199	131.05	
8		4	194	1.48248	130.86	
9	3	1	78	0.60661	128.58	
10		2	119	0.91892	129.50	
11		3	128	0.99199	129.03	
12		4	201	1.48248	135.58	
13	4	1	81	0.60661	133.53	
14		2	134	0.91892	145.82	
15		3	141	0.99199	142.14	
16		4	216	1.48248	145.70	

Calculating TR_t(step 8)

Period(t)	d=deseasonalize Y	tr=113.685+1.856t	estimate Y = TRXSN
1	118.69	115.54	70.09
2	119.71	117.40	107.88
3	117.94	119.25	118.30
4	116.02	121.11	179.54
5	125.29	122.97	74.59
6	121.88	124.82	114.70
7	131.05	126.68	125.66
8	130.86	128.53	190.55
9	128.58	130.39	79.09
10	129.50	132.25	121.52
11	129.03	134.10	133.03
12	135.58	135.96	201.55
13	133.53	137.81	83.60
14	145.82	139.67	128.34
15	142.14	141.53	140.39
16	145.70	143.38	212.56

Regression for Estimating TR_t

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.930985
R Square	0.866734
Adjusted R	0.857215
Standard E	3.585966
Observati	16

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regressio	1	1170.863	1170.863	91.05289	1.66E-07
Residual	14	180.0281	12.85915		
Total	15	1350.891			

	Coefficients	*Standard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	113.6851	1.880497	60.45481	2.47E-18	109.6518	117.7183	109.6518	117.7183
X Variable	1.855725	0.194476	9.542164	1.66E-07	1.438615	2.272835	1.438615	2.272835

Measuring Forecast Accuracy :

Let $e_t = Y_t - \hat{Y}_t$ be the errors of forecast.

1) Mean Squared Error

$$MSE = \sum_{t=1}^{n} e_t^2 / n$$

$$RMSE = \sqrt{MSE}$$

MSE = 11.93
RMSE = 3.45

2) Mean Absolute Deviation

$$MAD = \sum_{t=1}^{n} |e_t| / n$$

$$RMAD = \sqrt{MAD}$$

MAD = 2.89
RMAD = 1.70

Additive Decomposition Method

Additive Models

$$y_t = TR_t + SN_t + CL_t + IR_t$$

Where.

y_t = the observed value of the time series at time t

TR_t = the trend component (or factor) at time t

SN_t = the seasonal component (or factor) at time t

CL_t = the cyclical component (or factor) at time t

IR_t = the irregular component (or factor) at time t

Additive Decomposition Method

Step by step approach

1) Calculate **12** month moving average (MA_t) - to eliminate **seasonal** & irregular fluctuation --- Question : If it was quarterly data what would you do???

2) Next calculate Centered Moving Average CMA_t :Why??

Clue : we are calculating moving average of even order. If it was of odd order, this step was not necessary

Since the model is $y_t = TR_t + SN_t + CL_t + IR_t$, this implies that

$$SN_t + IR_t = y_t - (TR_t + CL_t) = y_t - CMA_t$$

Because CMA_t is considered to be an estimate of $TR_t + CL_t$, as averaging process is assumed to have removed seasonal and short term irregular fluctuations

3) Calculate $SN_t + IR_t$

Additive Decomposition Method

Step by step approach

4) Next step calculate average SN_t for each season (Jan, Feb, ...)

5) Then do normalization of average SN_t . As L is the number of seasons the normalization is achieved by subtracting $\sum_1^L SN_t$/L from average SN_t so that the sum of the normalized $SN_t = 0$

6) Next step is to calculate the de-seasonalized observation d_t at time t:

$$d_t = y_t - \text{ estimated } SN_t$$

Deseasonalized observations are computed in order to better estimate the trend components. By subtracting seasonality from y_t we can have better understanding of trend

7) Next step is to estimate TR_t . Look at plot of d_t , if it is linear please fit $TR_t = \beta_0 + \beta_1 t$ or if quadratic fit $TR_t = \beta_0 + \beta_1 t + \beta_2 t^2$

This completes the estimation of SN_t (step 6) and TR_t

Additive Decomposition Method

Step by step approach

8) So from the equation $y_t = TR_t + SN_t + CL_t + IR_t$

it implies that $CL_t + IR_t = y_t - (TR_t + SN_t)$ where one can use the estimate of TR_t and SN_t

9) It is observed empirically that when considering monthly or quarterly data we can average out ir_t by taking a three period moving average

10) Finally we calculate the estimate of IR_t by using the formula,

$$(cl_t + ir_t) - cl_t$$

Note : Traditionally the estimates tr_t, sn_t, cl_t and ir_t are obtained by using additive decomposition method and used to describe the time series.

However we can also use these estimates to forecast the future values of time series. If there is no pattern in irregular component, we predict IR_t to be 0.

Additive Decomposition Method

Forecasting using the method: In case $IR_t = 0$,

A) The point forecast of $y_t = tr_t + sn_t + cl_t$, if a well defined cycle exists and can be predicted

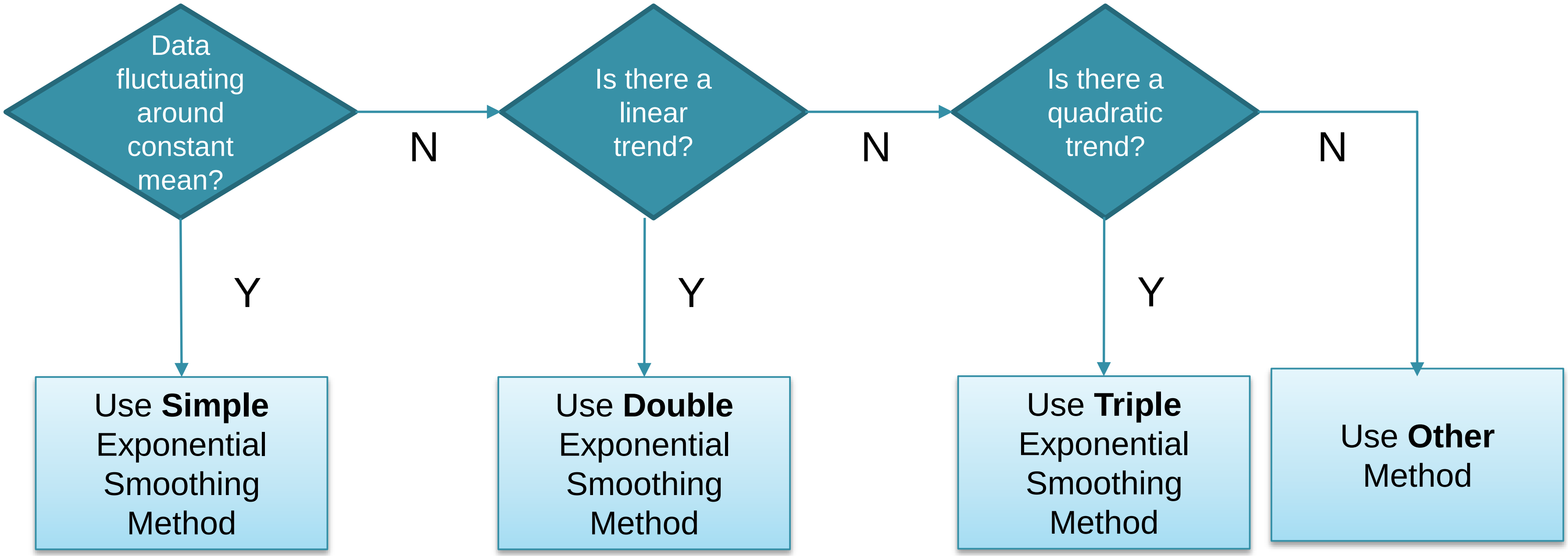
B) The point forecast of $y_t = tr_t + sn_t$, if a well defined cycle does not exist or if CL_t can not be predicted

Although there is no theoretically correct prediction interval for y_t Bowerman , O'Connell & Koehler have found that a fairly accurate (approx.) $100(1-\alpha)\%$ prediction interval for y_t is

$[y_t \pm B_t[100(1-\alpha)]]$ where $B_t[100(1-\alpha)]$ is a error bound in a $100(1-\alpha)\%$ prediction interval $[tr_t \pm B_t[100(1-\alpha)]]$ for the deseasonalized observation

$$d_t = y_t - sn_t$$

Three types of Exponential smoothing methods



Revisit Smoothing Methods

Smoothing

- Respond to the most recent behavior of the series
- Employ the idea of weighted averages
- They range in the degree of sophistication

1. Simple Averages - quick, inexpensive (should only be used on stationary data)
2. Moving Averages - a constant number specified at the outset and a mean computed for the *most recent observations* - such as a 3 or 4 period moving average.
    - Works best with stationary data.
    - The larger the order of the moving average, the greater the smoothing effect. Larger *n* when there are wide, infrequent fluctuations in the data.
    - By smoothing recent actual values, removes randomness.

Smoothing Methods - Formula

- Simple average
  - m ~ most recently observation

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + .. + Y_{t-m+1}}{m}$$

- Moving average Smoothing (Trailing moving average)
  - w ~ window / width or interval

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + .. + Y_{t-w+1}}{w}$$

- Simple Exponential Smoothing
  - Smoothing parameter $0 < \alpha < 1$

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

Exponential Smoothing

- Exponential Smoothing provides a forecasting method that is most effective when the components (trend and seasonal factors) of the time series may be changing over time
- More recent observations are weighted more heavily than more remote observations
- The unequal weighting is accomplished by using one or more smoothing constants, which determine how much weight is given to each observation

Exponential Smoothing

- If $y(1), y(2), \ldots$, is a time series of observations in periods 1, 2, … and $F(2), F(3), \ldots$ the series of forecasts in periods 2, 3, … then

$$F(t+1) = \alpha y(t) + (1-\alpha)F(t)$$
$$= \alpha y(t) + (1-\alpha)\{\alpha\, y(t-1) + (1-\alpha)F(t-1)\}$$
$$= \alpha y(t) + \alpha(1-\alpha)y(t-1) + (1-\alpha)\,\{\alpha\, y(t-2) + (1-\alpha)F(t-2)\}$$
$$= \alpha y(t) + \alpha(1-\alpha)y(t-1) + \alpha(1-\alpha)\,\{y(t-2) + \ldots$$

- Underlying assumption – the time series has a constant or slowly changing mean
  - No trend or seasonality
- How large should α be?
  - If $\alpha = 1$, the forecast is the last observed value
  - If $\alpha = 0$, the forecast is the last forecast
  - Large α (say 0.8) tends to capture changes in the underlying process fast but are sensitive to noise (random fluctuations)
  - A compromise is required; values between 0.1 and 0.25 often used

Revisit Simple Exponential Method

- Data fluctuating around a constant mean, there is no trend .
- Consider **no trend equation**, where β_0 is slowly changing over time.

$$Y_t = \beta_0 + \varepsilon_t$$

Step 1 : Find Initial estimate $a_0(0)$

$$a_{0(0)} = \bar{y} = \frac{\sum y_t}{n}$$

Step 2 : Compute $a_0(T)$ with an arbitrary α

$$a_{0(T)} = \alpha y_T + (1-\alpha) a_{0(T-1)}$$

Step 3 : Compute MSE/RMSE/MAD/MAPE

$$\text{MSE} = \sum_{i=1}^{n} \frac{\left(Y_i - \hat{Y}_i\right)^2}{n}$$

Step 4 : Select α that yield the lowest in step 3.

Holt's Trend Corrected Exponential Smoothing (2 parameters)

Suppose a time series displays a linear trend. Then the series can be described by the linear trend model :

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

Therefore level (or mean) at time T is $\beta_0 + \beta_1$ T and

level (or mean) at time T-1 is $\beta_0 + \beta_1$ (T-1)

So increase or decrease in the level of time series from time T-1 to T is:

$$(\beta_0 + \beta_1 T) - (\beta_0 + \beta_1 (T-1)) = \beta_1$$

This fixed rate of increase or decrease β_1 is called growth rate

Holt's trend corrected exponential smoothing is appropriate when both the level and the growth rate are changing

Holt's Trend Corrected Exponential Smoothing (2 parameters)

A model different from the linear model is needed to describe the changing level and growth rate

To implement Holt's trend corrected exponential smoothing we let ℓ_{T-1} denote the estimate of the level of the time series at time period T-1 and we denote b_{T-1} denote the estimate of the growth rate of the time series in time T-1. Then if we observe a new time series value y_T in time period T, we use two smoothing equations to update the estimates ℓ_{T-1} and b_{T-1}

$$\ell_T = \alpha y_T + (1-\alpha)(\ell_{T-1} + b_{T-1})$$

$$b_T = \gamma[\ell_T - \ell_{T-1}] + (1-\gamma) b_{T-1}$$

where α and γ are smoothing constants between 0 and 1

Additive Holt-Winters Method (3 parameters)

Suppose a time series displays a linear trend locally and has a seasonal pattern with constant (additive) seasonal variation and that the level, growth rate and seasonal pattern may be changing. Then the estimate ℓ_T for the level, the estimate b_T for the growth rate, and the estimate sn_T for the seasonal factor of the time series in time T is given by the smoothing equations :

$$\ell_T = \alpha (y_T - sn_{T-L}) + (1-\alpha)(\ell_{T-1} + b_{T-1})$$

$$b_T = \gamma [\ell_T - \ell_{T-1}] + (1-\gamma) b_{T-1}$$

$$sn_T = \delta [y_T - \ell_T] + (1-\delta) sn_{T-1}$$

where α , γ and δ are smoothing constants between 0 and 1

Workshop