FIT3158
Business Decision Modelling

Lecture 10
Forecasting

Topics Covered:

- The Components of a Time Series
- Moving Average Forecasts
- Measures of Forecast Accuracy
- Simple Exponential Smoothing
- Forecasting Using Linear Regression
- Seasonal forecasting: Multiplicative & Additive Model

What is a Time Series?
What is Forecasting?

- A time series is a set of data observed over time.
- Forecasting is the practice of predicting future values.

Quantitative Approaches to Forecasting

- Quantitative methods are based on an analysis of historical data concerning one or more time series.
- If the historical data used are restricted to past values of the series that we are trying to forecast, the procedure is called a time series method.
- If the historical data used involve other time series that are believed to be related to the time series that we are trying to forecast, the procedure is called a causal method.
- Three time series methods are: smoothing, trend projection and trend projection adjusted for seasonal influence

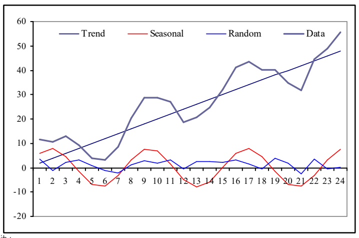
Components of a Time Series

- The trend component accounts for the gradual shifting of the time series over a long period of time.
- Any regular pattern of sequences of values above and below the trend line might be attributable to the cyclical component of the series.
- The seasonal component of the series accounts for regular patterns of variability within certain time periods, such as over a year.
- The irregular component of the series is caused by (short-term) unanticipated and non-recurring factors that affect the values of the time series. One cannot attempt to predict its impact on the time series in advance.
- These components may be put together to form either *additive* or *multiplicative* (and potentially both) time series.

Additive Model

- Data can be fitted by an additive model when it is reasonable to assume that the observed time series can be explained as:

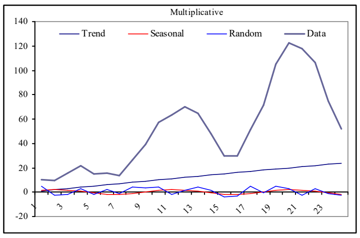
Data = Trend + Seasonal Variation + Random Variation



Multiplicative Model

- Data can be fitted by a multiplicative model when it is reasonable to assume that the observed time series can be explained as:

*Data = Trend * Seasonal Variation * Random Variation*



Some Time Series Terms

- **Stationary Data**:

  A time series variable exhibiting no significant upward or downward trend over time, and no change in variability over time.

- **Non-stationary Data**:

  A time series variable exhibiting (e.g.) a significant upward or downward trend over time or (e.g.) changing variability over time.

- **Seasonal Data**:

  A time series variable exhibiting a repeating patterns at regular intervals over time (e.g., hours, weeks, months, years, etc.).

Extrapolation Models

- Extrapolation models try to account for past behavior of a time series variable in an attempt to predict the future behavior of the variable

Actual value for period t

Actual value for period t-2

$$\hat{Y}_{t+1} = f\left(Y_t, Y_{t-1}, Y_{t-2}, \ldots\right)$$

Predicted value in time period t+1

Actual value for period t-1

- We will look at several extrapolation techniques that are appropriate for stationary data: smoothing methods including moving averages and exponential smoothing.

Smoothing Methods

- In cases in which the time series is fairly stable and has no significant trend, seasonal, or cyclical effects, one can use smoothing methods to average out the irregular components of the time series.
- Four common smoothing methods, which are particularly useful for additive models, are:
  - **Moving averages**
  - **Centred moving averages**
  - **Weighted moving averages**
  - **Exponential smoothing**

Moving Average Forecasting

- *k > 0* is an integer.
- No (widely accepted) general method exists for determining *k*.
- Try out several *k* values to (somehow) see what works best.

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + \ldots + Y_{t-k+1}}{k}$$

How many previous observations will be included?

The moving average method (MA method) consists of computing an average of the most recent *n* data values for the series and using this average for forecasting the value of the time series for the next period.

Example: VCR Data

- VCR sales have been recorded over 24 months.
- Construct 2-Month and 4-Month Moving Averages (MAs) of the data.
- Forecast the expected number sold during months 25 to 28.

See File : Lecture 10.xlsx

(VCRs 2MA 4MA)

Time Period	Number of VCRs Sold	2-Month Moving Avg	4-Month Moving Avg
1	33	--	--
2	38	--	--
3	31	35.50	--
4	35	34.50	--
5	30	33.00	34.25
6	36	32.50	33.50
7	34	33.00	33.00
8	39	35.00	33.75
9	39	36.50	34.75
10	36	39.00	37.00
11	40	37.50	37.00
12	38	38.00	38.50
13	37	39.00	38.25
14	39	37.50	37.75
15	32	38.00	38.50
16	38	35.50	36.50
17	37	35.00	36.50
18	39	37.50	36.50
19	37	38.00	36.50
20	35	38.00	37.75
21	37	36.00	37.00
22	34	36.00	37.00
23	35	35.50	35.75
24	36	34.50	35.25
25		35.50	35.50
26		35.75	35.13
27		35.63	35.41
28		35.69	35.51

Example: VCR Data

- For the two month moving average:

$$\hat{Y}_{25} = \frac{Y_{24} + Y_{23}}{2} = \frac{36 + 35}{2} = 35.5$$

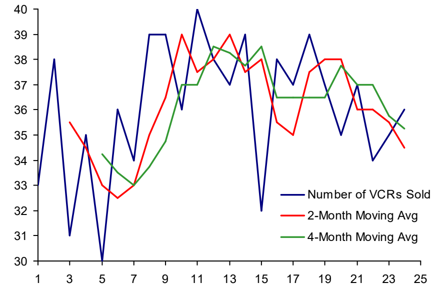
$$\hat{Y}_{26} = \frac{\hat{Y}_{25} + Y_{24}}{2} = \frac{35.5 + 36}{2} = 35.75$$

- For the four month moving average:

$$\hat{Y}_{25} = \frac{Y_{24} + Y_{23} + Y_{22} + Y_{21}}{4} = \frac{36 + 35 + 34 + 37}{4} = 35.5$$

$$\hat{Y}_{26} = \frac{\hat{Y}_{25} + Y_{24} + Y_{23} + Y_{22}}{4} = \frac{35.5 + 36 + 35 + 34}{4} = 35.13$$

Example: VCR Data



Use of Moving Averages in Share Trading

One approach to determine when to buy a particular share is to plot the 9 day and 21 day moving average forecasts. The points where these cross is used by some traders to signal buy and sell points. United Group daily from E*Trade:

Weighted moving average

- The moving average technique assigns equal weight to all previous observations:

$$\hat{Y}_{t+1} = \frac{1}{k}Y_t + \frac{1}{k}Y_{t-1} + \cdots + \frac{1}{k}Y_{t-k+1}$$

- The **weighted moving average** technique allows for different weights to be assigned to previous observations:

$$\hat{Y}_{t+1} = w_1 Y_t + w_2 Y_{t-1} + \cdots + w_k Y_{t-k+1}$$

- We must determine values for k and the w_i.
- We usually determine these to minimise error – measured in several different ways.

Weighted moving average

Using the Solver to determine the optimal weights, we see that w_1=.291 and w_2=.709 minimises Mean Squared Error (MSE).

Time Period	Number of VCRs Sold	2-Month Weighted Moving Avg		Weights	
1	33	--		w1	0.291
2	38	--		w2	0.709
3	31	34.46		sum	1.000
4	35	35.96			
5	30	32.17		MSE	6.29
6	36	33.54			
7	34	31.75			
8	39	35.42			
9	39	35.46			
10	36	39.00			
11	40	38.13			
12	38	37.17			
13	37	39.42			
14	39	37.71			
15	32	37.58			
16	38	36.96			
17	37	33.75			
18	39	37.71			
19	37	37.58			
20	35	38.42			
21	37	36.42			
22	34	35.58			
23	35	36.13			
24	36	34.29			
25		35.29			
26		35.79			
27		35.44			
28		35.69			

See File : Lecture 10.xlsx

(VCRs Weighted MA)

Forecast Accuracy

- We can determine the accuracy of the forecast values by comparing them with actual (subsequent) observations. *Note that we don't compare the fitted model with the observations used to construct the model.*
- There are a variety techniques to determine the forecast accuracy:
  - Mean Squared Error
  - Mean Absolute Percent Error
  - Mean Absolute Deviation
  - Correlation Coefficient (**When we can assume a linear model.*)

Forecast Accuracy

These are some measures of forecast accuracy:

- Mean Squared Error (MSE) : The average of the squared forecast errors for the historical data is calculated. The forecasting method or parameter(s) which minimize this mean squared error is then selected.
- Mean Absolute Deviation (MAD, sometimes also called mean absolute error) : The mean of the absolute values of all forecast errors is calculated. The mean absolute deviation measure is less sensitive to individual large forecast errors than the mean squared error measure.
- Mean Absolute Percent Error (MAPE) : The absolute value of all forecast errors are divided by the observed value. The mean is then calculated. This method scales each error against the true value.

MSE, MAD, MAPE

$$MSE = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n} \quad MAD = \frac{\sum_{i=1}^{n}|Y_i - \hat{Y}_i|}{n} \quad MAPE = \frac{\sum_{i=1}^{n}\frac{|Y_i - \hat{Y}_i|}{Y_i}}{n}$$

Y_i are the actual (observed) values

$\hat{Y}_i$ are the fitted (forecast) values

n is the number of forecast values

Adaptive Forecasting

- **Adaptive forecasting** methods are used to make short-term forecasts. These methods may be appropriate when no constant long-term trend is present in the data.
- The term 'adaptive' refers to the fact that each period's forecast is updated by subsequent observations before the next forecast is made.
- Adaptive methods include:
  - Simple Exponential Smoothing to predict one period ahead *(one-step forecasting)*
  - Double Exponential Smoothing - to include a growth or trend factor, and to predict one, two ... periods into the future.

Simple Exponential Smoothing

- Using exponential smoothing, the forecast for the next period is equal to the forecast for the current period plus a proportion (α) of the forecast error in the current period.

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

where

$\hat{Y}_t$ is the forecast value

Y_t is the observed value

α is the smoothing factor, $0 < \alpha < 1$

t is the time index

Simple Exponential Smoothing cont…

- Think of the model as:

New forecast = current forecast + α(error)

Where: α is between 0 and 1

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

Forecast Next Period, Forecast Current Period, Observed Current Period

Example: Sale of Headache Tablets

- During the past ten weeks, sales of cases of a headache medicine have been as follows:
- If we use exponential smoothing to forecast sales, which value for the smoothing constant α, α = .1 or α = 0.8 (or possibly something else), gives better forecasts?

See File : Lecture 10.xlsx

(Headache Exp)

Week (t)	Sales (t)
1	110.0
2	115.0
3	125.0
4	120.0
5	125.0
6	120.0
7	130.0
8	115.0
9	110.0
10	130.0

Example:

Exponential Smoothing

- To evaluate the two smoothing constants, determine how the forecasted values would compare with the actual historical values in each case.

Let: Y_t = actual sales in week t

$\hat{Y}_t$ = forecasted sales in week t

$\hat{Y}_1 = Y_1 = 110$

For other weeks, $\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)\hat{Y}_t$

Example:

Exponential Smoothing: For $\alpha = 0.1$, $1 - \alpha = 0.9$

$\hat{Y}_1 = 110$

$\hat{Y}_2 = .1Y_1 + .9\hat{Y}_1 = .1(110) + .9(110) = 110$

$\hat{Y}_3 = .1Y_2 + .9\hat{Y}_2 = .1(115) + .9(110) = 110.5$

$\hat{Y}_4 = .1Y_3 + .9\hat{Y}_3 = .1(125) + .9(110.5) = 111.95$

$\hat{Y}_5 = .1Y_4 + .9\hat{Y}_4 = .1(120) + .9(111.95) = 112.76$

$\hat{Y}_6 = .1Y_5 + .9\hat{Y}_5 = .1(125) + .9(112.76) = 113.98$

$\hat{Y}_7 = .1Y_6 + .9\hat{Y}_6 = .1(120) + .9(113.98) = 114.58$

$\hat{Y}_8 = .1Y_7 + .9\hat{Y}_7 = .1(130) + .9(114.58) = 116.12$

$\hat{Y}_9 = .1Y_8 + .9\hat{Y}_8 = .1(115) + .9(116.12) = 116.01$

$\hat{Y}_{10} = .1Y_9 + .9\hat{Y}_9 = .1(110) + .9(116.01) = 115.41$

Modelling it in the Spreadsheet

See File : Lecture 10.xlsx
(Headache Exp)

alpha =	0.1	
Week (t)	Sales (t)	forecast
1	110.0	110.00
2	115.0	110.00
3	125.0	110.50
4	120.0	111.95
5	125.0	112.76
6	120.0	113.98
7	130.0	114.58
8	115.0	116.12
9	110.0	116.01
10	130.0	115.41
11		116.87

alpha =	0.1	
Week (t)	Sales (t)	forecast
1	110	=B4
2	115	=$B$1*B4+(1-$B$1)*C4
3	125	=$B$1*B5+(1-$B$1)*C5
4	120	=$B$1*B6+(1-$B$1)*C6
5	125	=$B$1*B7+(1-$B$1)*C7
6	120	=$B$1*B8+(1-$B$1)*C8
7	130	=$B$1*B9+(1-$B$1)*C9
8	115	=$B$1*B10+(1-$B$1)*C10
9	110	=$B$1*B11+(1-$B$1)*C11
10	130	=$B$1*B12+(1-$B$1)*C12
11		=$B$1*B13+(1-$B$1)*C13

Comparison of performance measures:

- when $\alpha = 0.1$.

alpha =	0.1				
Week (t)	Sales (t)	forecast	APE	AD	SE
1	110.0	110.00	-	-	-
2	115.0	110.00	0.04	5.00	25.00
3	125.0	110.50	0.12	14.50	210.25
4	120.0	111.95	0.07	8.05	64.80
5	125.0	112.76	0.10	12.25	149.94
6	120.0	113.98	0.05	6.02	36.25
7	130.0	114.58	0.12	15.42	237.73
8	115.0	116.12	0.01	1.12	1.26
9	110.0	116.01	0.05	6.01	36.13
10	130.0	115.41	0.11	14.59	212.87
			MAPE	MAD	MSE
			0.07	9.22	108.25

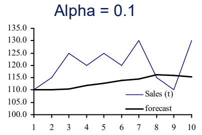
Comparison of performance measures:

- when $\alpha = 0.2$.

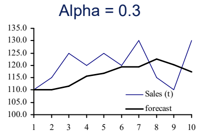
alpha =	0.2				
Week (t)	Sales (t)	forecast	APE	AD	SE
1	110.0	110.00	-	-	-
2	115.0	110.00	0.04	5.00	25.00
3	125.0	111.00	0.11	14.00	196.00
4	120.0	113.80	0.05	6.20	38.44
5	125.0	115.04	0.08	9.96	99.20
6	120.0	117.03	0.02	2.97	8.81
7	130.0	117.63	0.10	12.37	153.13
8	115.0	120.10	0.04	5.10	26.01
9	110.0	119.08	0.08	9.08	82.45
10	130.0	117.26	0.10	12.74	162.20
			MAPE	MAD	MSE
			0.07	8.60	87.92

The Value of alpha

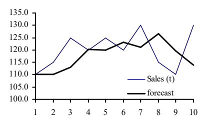
Alpha = 0.1



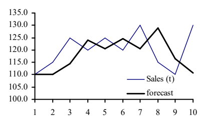
Alpha = 0.3



Alpha = 0.6



Alpha = 0.9



The effect of α on forecast accuracy

The table below shows each measure of forecast accuracy as alpha varies.

Alpha	APE	MAD	MSE
0.00	0.09	11.11	166.67
0.10	0.07	9.22	108.25
0.20	0.07	8.60	87.92
0.30	0.07	8.11	81.31
0.40	0.07	7.94	80.21
0.50	0.07	7.91	81.81
0.60	0.06	7.91	84.98
0.70	0.07	8.17	89.21
0.80	0.07	8.42	94.17
0.90	0.07	8.66	99.65
1.00	0.07	8.89	105.56

Using Solver to Determine α

- Instead of guessing the optimal value of alpha it is possible to use the solver to determine the optimal value of alpha.
- Lowest MSE of *12.45* is obtained when *a = 0.9.*
- *Note that the forecast is (arguably) only meaningful for period 25. Why?*

See File : Lecture 10.xlsx
(VCR E Smooth)

Time Period	Number of VCRs Sold	Exp. Smoothing Prediction			
1	33	33.00		alpha	0.900
2	38	33.00			
3	31	37.50		MSE	12.45
4	35	31.65			
5	30	34.67			
6	36	30.47			
7	34	35.45			
8	39	34.14			
9	39	38.51			
10	36	38.95			
11	40	36.30			
12	38	39.63			
13	37	38.16			
14	39	37.12			
15	32	38.81			
16	38	32.68			
17	37	37.47			
18	39	37.05			
19	37	38.80			
20	35	37.18			
21	37	35.22			
22	34	36.82			
23	35	34.28			
24	36	34.93			
25		35.89			
26		35.89			
27		35.89			
28		35.89			

Trend Projection

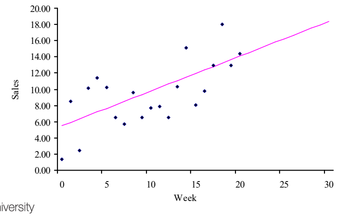
- If a time series exhibits a linear trend, the method of least squares may be used to determine a trend line (projection) for future forecasts.
- Least squares, also used in regression analysis, determines the unique trend line forecast which minimizes the mean square error between the trend line forecasts and the actual observed values for the time series. The method of least squares will result from – and can be justified by – assuming a Normal distribution and then estimating using Maximum Likelihood estimation (MLE)
  - although not everyone is convinced about the merits of Maximum Likelihood estimation
- The independent variable is the time period and the dependent variable is the actual observed value in the time series.

Linear Regression

- Regression is the practice of describing the relationship between 2 or more quantitative variables. Thus, if we know the value of one variable, we can estimate the value of the related variable of interest.
- We may use this to develop a model (or equation) of our data.
- We may use the model for prediction or extrapolation.
- Regression based methods are particularly suited to long-term forecasting.
- For example:
  - The time it takes to serve a customer given the number of items they have purchased
  - The industry standard for the number of staff employed given the annual turnover of the company
  - The future value of time series data.

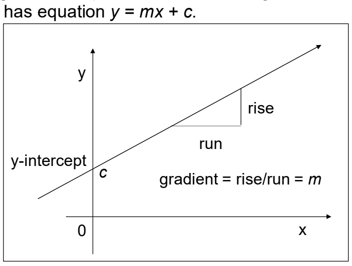
Regression-Based Forecasting

- The data should be linear and non-seasonal. A straight line with the general equation *y = ax + b* is fitted to the data. The equation of this line then forms the basis for prediction.



The Equation of a Straight Line

- We can use the basic equation of a straight line as the model for our regression equation. A line with gradient *m* and y-intercept *c* has equation *y = mx + c*.



Least Squares Regression

- The basic idea is to find the straight line which minimises the squared differences between actual points and those predicted by the model.
- To express the regression of y on x as $y = mx + c$, we calculate:

$$m = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} \quad \text{and} \quad c = \bar{y} - m\bar{x}$$

Or

$$m = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

Forecasting Real Data

- Forecast the demand for Ready-Mixed Concrete for 1999 - 2001 using the known demand for concrete over the years 1976 - 1998.
- We will assume that these data do not have a seasonal pattern.
- Before we fit our model we need to *index* the data by numbering each observation with a period number.
- We then determine the intercept and slope using the built-in functions in Excel or the formula (on approx. slide 37).

See File : Lecture 10.xlsx

(Concrete Regression)

Date	Ready mixed concrete ('000 cubic metres)
Jun.1976	2700
Sep.1976	2790
Dec.1976	2652
Mar.1977	2210
Jun.1977	2592
Sep.1977	2677
Dec.1977	2724
Mar.1978	2282
Jun.1978	2694
Sep.1978	2793
Dec.1978	2841
Mar.1979	2543
Jun.1979	2661
Sep.1979	2828
Dec.1979	2820
Mar.1980	2645

Regression in EXCEL

- Regression is one of the analysis functions in EXCEL. However, you can also calculate the formulas manually with the following built-in formulas:
- *m* = SLOPE(y values, x values)
- *c* = INTERCEPT(y values, x values)
- The correlation coefficient measures the strength and the direction of a linear relationship between two variables, $r = \frac{s_{xy}}{s_x s_y}$

  Where s_x and s_y are the sample standard deviation of x and y,

  s_{xy} is the sample covariance of x and y.

  And so: $r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}}$
- The coefficient of determination, r^2, gives the proportion of the variation of one variable that is predictable from the other variable.
- r^2 = CORREL(y values, x values)^2

The Fitted Model

Date	Period	Ready mixed concrete ('000 cubic metres)	Predicted Ready Mixed Concrete
Jun.1976	1	2700	2608
Sep.1976	2	2790	2627
Dec.1976	3	2652	2646
Mar.1977	4	2210	2665
Jun.1977	5	2592	2683
Sep.1977	6	2677	2702
Dec.1977	7	2724	2721
Mar.1978	8	2282	2739
Jun.1978	9	2694	2758
Sep.1978	10	2793	2777
Dec.1978	11	2841	2795
Mar.1979	12	2543	2814
Jun.1979	13	2661	2833
Sep.1979	14	2828	2852
Dec.1979	15	2820	2870
Mar.1980	16	2645	2889

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.790224
R Square	0.624454
Adjusted R Square	0.620235
Standard Error	385.3951
Observations	91

ANOVA

	df
Regression	1
Residual	89
Total	90

	Coefficients
Intercept	2589.689
Period	18.71005

*predicted = period * m + c*

The Forecast

- We continue the predicted values past the model data in order to make the forecast.

Date	Period	Ready mixed concrete ('000 cubic metres)	Predicted Ready Mixed Concrete
Mar.1999	92		4311
Jun.1999	93		4919
Sep.1999	94		4947
Dec.1999	95		4974
Mar.2000	96		5001
Jun.2000	97		5028
Sep.2000	98		5055
Dec.2000	99		5082
Mar.2001	100		5110
Jun.2001	101		5137
Sep.2001	102		5164
Dec.2001	103		5191

Actual Data/Forecast Accuracy

y = mx + c
= 18.71 * 92 + 2589.69
= 4311

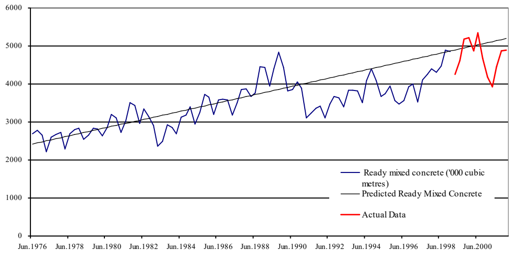
Date	Period	Predicted Ready Mixed Concrete	Actual Data	APE	Squared Errors
Mar.1999	92	4311	4251	0.01	3601.67
Jun.1999	93	4919	4611	0.07	95086.35
Sep.1999	94	4947	5187	0.05	57828.48
Dec.1999	95	4974	5217	0.05	59200.64
Mar.2000	96	5001	4880	0.02	14605.18
Jun.2000	97	5028	5349	0.06	103030.99
Sep.2000	98	5055	4679	0.08	141510.87
Dec.2000	99	5082	4188	0.21	799849.43
Mar.2001	100	5110	3932	0.30	1386522.09
Jun.2001	101	5137	4452	0.15	468773.61
Sep.2001	102	5164	4881	0.06	79995.16
Dec.2001	103	5191	4896	0.06	87023.74
			R-Squared	MAPE	MSE
			0.03	0.09	274752.35

Having observed these values (Mar 1999 onwards) after making the forecast (from data up to 1998), we can now determine the forecast accuracy.

Some Observations:

- The model: $y = 18.71*t + 2589.69$.
- Thus, cement production increases by approximately 18,700 cubic metres each quarter, from a base level of approx 2,600,000 cubic metres in 1976.
- The MAPE of 9% indicates that the model predicts to a reasonable level of accuracy. However, the very low R-squared (or R^2) contradicts this and indicates a poor fit. This could be further investigated.
- The plotted data and model show a reasonable level of fit.

Data, Model and Forecast



Linear Regression - Seasonal Data

- There are two main approaches to regression-based seasonal forecasting.
  1. Additive seasonal factors - the effect of the seasons is due to a constant fluctuation through time.
  2. Multiplicative seasonal factors - the effect of the seasons is due to increasing or decreasing fluctuations through time

- Note that when data is seasonal, one of the conditions for simple least squares regression (that data does not have any systematic fluctuations) is violated. In order to use regression, we either (e.g.) de-seasonalise the data (multiplicative models) or (e.g.) introduce additional variables to account for the seasonal effects (additive models).

Multiplicative Time Series Model

- Also called the `Ratio-to-Moving-Average Method'
- Steps
  1. Calculate the centred moving averages (CMAs).
  2. Centre the CMAs on integer-valued periods.
  3. Determine the seasonal and irregular factors ($S_t I_t$).
  4. Determine the average seasonal factors.
  5. Scale the seasonal factors (S_t).
  6. Determine the de-seasonalised data.
  7. Determine a trend line of the de-seasonalised data.
  8. Determine the de-seasonalised predictions.
  9. Re-adjust forecast to account the seasonality.

Example: Terry's Tie Shop

- Business at Terry's Tie Shop can be viewed as falling into three distinct seasons: (1) Christmas (November-December); (2) Father's Day (late May - mid-June); and (3) all other times. Average weekly sales (in $'s) during each of these three seasons during the past four years has been as follows:

			Year	
Season	1	2	3	4
1	1856	1995	2241	2280
2	2012	2168	2306	2408
3	985	1072	1105	1120

- Determine a forecast for the average weekly sales in year 5 for each of the three seasons.

Example: Terry's Tie Shop

Spreadsheet showing data, seasonal indices and de-seasonalised data.

Year	Season	Period	Sales Y_t	Moving Average	Seasonal Index $S_t I_t$	Scaled S_t	Y_t/S_t
1	1	1	1856			1.178	1576
	2	2	2012	1617.67	1.244	1.236	1628
	3	3	985	1664.00	0.592	0.586	1681
2	1	4	1995	1716.00	1.163	1.178	1694
	2	5	2168	1745.00	1.242	1.236	1754
	3	6	1072	1827.00	0.587	0.586	1829
3	1	7	2241	1873.00	1.196	1.178	1902
	2	8	2306	1884.00	1.224	1.236	1866
	3	9	1105	1897.00	0.582	0.586	1886
4	1	10	2280	1931.00	1.181	1.178	1935
	2	11	2408	1936.00	1.244	1.236	1948
	3	12	1120			0.586	1911

See File : Lecture 10.xlsx (Terry's Ties Multi Model)

Example: Terry's Tie Shop

1. Calculate the centred moving averages (CMAs).

   There are three distinct seasons in each year. Hence, take a three season moving average to eliminate seasonal and irregular factors. For example the first moving average is:

   (1856 + 2012 + 985)/3 =1617.67.

2. Centre the CMAs on integer-valued periods.

   The first moving average computed in step 1 (1617.67) will be centred on season 2 of year 1. Note that the moving averages (Mas) from step 1 centre themselves on integer-valued periods because *n* is an odd number.

Cont'd …

3. Determine the seasonal and irregular factors (S_t, I_t).

   Isolate the trend and cyclical components. For each period *t*, this is given by:

   Y_t / (Moving Average for period *t*).

4. Determine the average seasonal factors.

   Averaging all $S_t I_t$ values corresponding to that season:

   Season 1: (1.163 + 1.196 + 1.181) /3 = 1.180

   Season 2: (1.244 + 1.242 + 1.224 + 1.244) /4 = 1.238

   Season 3: (.592 + .587 + .582) /3 = 0.587

Cont'd …

5. Scale the seasonal factors (S_t).

   Divide each seasonal factor by the average of the seasonal factors.

   Average of the seasonal factors = (1.180 + 1.238 + .587)/3 = 1.002

   Season 1: 1.180/1.002 = 1.178

   Season 2: 1.238/1.002 = 1.236

   Season 3: .587/1.002 = 0.586

   Total = 3.000 (Average = 1.00)

6. Determine the de-seasonalised data.

   Divide the data point values, Y_t, by S_t.

Cont'd …

7. Determine a trend line of the deseasonalised data.

   Using the least squares method for *t* = 1, 2, ..., 12, gives: T_t = 1580.11 + 33.96*t*

- The trend line can be found using linear regression. When doing this manually we use the sum of squares table:

$$m = \frac{s_{xy}}{s_x^{\ 2}} = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

and $c = \bar{y} - m\bar{x}$

	X	Y	X^2	Y^2	XY
	1	1576	1	2483776	1576
	2	1628	4	2650384	3256
	3	1681	9	2825761	5043
	4	1694	16	2869636	6776
	5	1754	25	3076516	8770
	6	1829	36	3345241	10974
	7	1902	49	3617604	13314
	8	1866	64	3481956	14928
	9	1886	81	3556996	16974
	10	1935	100	3744225	19350
	11	1948	121	3794704	21428
	12	1911	144	3651921	22932
Sum	78	21610	650	39098720	145321

Alternatively, …

7. You can determine the trend line from the regression table:

   Ŷ = mx + c

   Ŷ = 33.96x + 1580.11

Y = mx + c
m = period = 33.958
c = Intercept = 1580.11

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.95001
R Square	0.90252
Adjusted R Square	0.89277
Standard Error	42.2036
Observations	12

ANOVA

	df	*SS*	*MS*	*F*
Regression	1	164900	164900.25	92.5812
Residual	10	17811.4	1781.1415	
Total	11	182712		

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	1580.11	25.9745	60.832966	3.5E-14
Period	33.958	3.52924	9.6219135	2.3E-06

Cont'd …

8. Determine the de-seasonalised predictions.

   Substitute *t* = 13, 14, and 15 into the least squares equation:

   T_{13} = 1580.11 + (33.96)(13) = 2022

   T_{14} = 1580.11 + (33.96)(14) = 2056

   T_{15} = 1580.11 + (33.96)(15) = 2090

9. Take into account the seasonality.

   Multiply each de-seasonalised prediction by its seasonal factor to give the following forecasts:

   Season 1: (1.178)(2022) = 2381

   Season 2: (1.236)(2056) = 2541

   Season 3: (.586)(2090) = 1224

Additive Time Series Model

- This method presumes that the data fits the following model:

*data = a * trend + b * seasonal effect + constant*

Steps:

1. Index the data by successive time periods.
2. Create a set of indicator variables 0 and 1 in {0,1}, where 1 indicates that the observation occurs in a certain season.
3. Determine the multiple regression of the observed data on the time period and seasonal indicators.

Example: Terry's Tie Shop

- Spreadsheet showing data, period number and seasonal indicator variables.

Year	Season	Period	S_1 Indicator	S_2 Indicator	S_3 Indicator	Sales Y_t
1	1	1	1	0	0	1856
	2	2	0	1	0	2012
	3	3	0	0	1	985
2	1	4	1	0	0	1995
	2	5	0	1	0	2168
	3	6	0	0	1	1072
3	1	7	1	0	0	2241
	2	8	0	1	0	2306
	3	9	0	0	1	1105
4	1	10	1	0	0	2280
	2	11	0	1	0	2408
	3	12	0	0	1	1120

See File : Lecture 10.xlsx (Terry's Ties Additive Model)

Fitting a Multiple Regression

The fitted model is thus:

Predicted =

797.00 + 36.47 * Period

+ 1095.43 * S1

+ 1189.47 * S2

+ 0 * S3

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.99
R Square	0.99
Adjusted R Square	0.86
Standard Error	73.09
Observations	12
	Coefficients
Intercept	797.00
Period	36.47
S1 Indicator	1095.43
S2 Indicator	1189.47
S3 Indicator	0.00

The Fitted Model

Year	Season	Period	S_1 Indicator	S_2 Indicator	S_3 Indicator	Fitted
1	1	1	1	0	0	1929
	2	2	0	1	0	2059
	3	3	0	0	1	906
2	1	4	1	0	0	2038
	2	5	0	1	0	2169
	3	6	0	0	1	1016
3	1	7	1	0	0	2148
	2	8	0	1	0	2278
	3	9	0	0	1	1125
4	1	10	1	0	0	2257
	2	11	0	1	0	2388
	3	12	0	0	1	1235
5	1	13	1	0	0	2367
	2	14	0	1	0	2497
	3	15	0	0	1	1344
Coefficients	797.00	36.47	1095.43	1189.47	0.00	

Centred Moving Average Smoothing

- The centred moving average method consists of computing an average of *n* periods' data and associating it with the midpoint of the periods. For example, the average for periods 5, 6, and 7 is associated with period 6.
- When the smoothing period is an even number, the moving average needs to be centred in order that the smoothed data coincides with an actual period.
- This methodology is a smoothing process, rather than a forecasting process. However, this method is useful in the process of computing seasonal indexes (or indices).

Example:

- Microsoft Excel Spreadsheet Showing moving averages when n = 3, 4 and 5. The MA4 has been centred (by averaging pairs).

Week (t)	Sales (t)	MA(3)	MA(5)	MA(4)	CMA(4)
1	110.0	-	-	-	-
2	115.0	116.7	-	117.5	-
3	125.0	120.0	119.0	121.3	119.4
4	120.0	123.3	121.0	122.5	121.9
5	125.0	121.7	124.0	123.8	123.1
6	120.0	125.0	122.0	122.5	123.1
7	130.0	121.7	120.0	118.8	120.6
8	115.0	118.3	121.0	121.3	120.0
9	110.0	118.3	-	-	-
10	130.0	-	-	-	-

See File : Lecture 10.xlsx (Headache CMA)

End of Lecture 10

References:

Ragsdale, C. (2021). Spreadsheet Modeling & Decision Analysis: A Practical Introduction to Business Analytics (9e) Cengage Learning: Chapters 9 & 11

Ragsdale, C. (2017). Spreadsheet Modeling & Decision Analysis: A Practical Introduction to Business Analytics (8e) Cengage Learning: Chapters 9 & 11

Lapin, L. and Whisler, W., Quantitative Decision Making with Spreadsheet Applications 7th Ed., Wadsworth (Thomson Learning) Belmont, 2002: Chapter 7

Tutorial 9 this week:

- Non-probabilistic methods – Maximax, Maximin and Minimax Regret decision rules.
- Probabilistic methods – EMV (Expected Monetary Value) and EOL (Expected Opportunity Loss) decision rules & the expected value of perfect information (EVPI)
- Solving decision problems using decision trees

Homework

- Familiarise yourself with the following:
  - ✓ Moving Average Forecasts
  - ✓ Simple Exponential Smoothing
  - ✓ Measures of Forecast Accuracy
  - ✓ Forecasting:
    - Using Linear Regression
    - Seasonal forecasting: Multiplicative & Additive Model

Readings for Lecture 11:

- Ragsdale, C. (2021). Spreadsheet Modeling & Decision Analysis: A Practical Introduction to Business Analytics (9e), Cengage Learning: Chapter 13
- Ragsdale, C. (2017). Spreadsheet Modeling & Decision Analysis: A Practical Introduction to Business Analytics (8e), Cengage Learning: Chapter 13