Forecasting

- Why Forecasting?
- How Do We Forecast?
  - Qualitative approach
    - based on experience, judgment, and knowledge
  - Quantitative approach
    - based on historical data and models
    - assume past patterns will continue into the future

Qualitative Forecasting
(based on experience, judgement, and knowledge)

- Grass Roots
- Market Research
- Panel Consensus
- Delphi Method
Quantitative Forecasting
(based on data and models)

• Casual Models:
  - Price
  - Population
  - Advertising
  ...→ Causal Model → Year 2009 Sales

• Time Series Models:
  - Sales_{2008}
  - Sales_{2007}
  - Sales_{2006}
  ...→ Time Series Model → Year 2009 Sales

Evaluation of Forecasting Model

• BIAS - The arithmetic mean of the errors
  \[
  BIAS = \frac{\sum (Actual - Forecast)}{n} = \frac{\sum Error}{n}
  \]
  - n is the number of forecast errors
  - Excel: =AVERAGE(error range)

• Mean Absolute Deviation - MAD
  \[
  MAD = \frac{\sum |Actual - Forecast|}{n} = \frac{\sum |Error|}{n}
  \]
  - No direct Excel function to calculate MAD

Evaluation of Forecasting Model

• Mean Square Error - MSE
  \[
  MSE = \frac{\sum (Actual - Forecast)^2}{n} = \frac{\sum (Error)^2}{n}
  \]
  - Excel: =SUMSQ(error range)/COUNT(error range)

• Mean Absolute Percentage Error - MAPE
  \[
  MAPE = \frac{\sum |Actual - Forecast|}{Actual} * 100\%
  \]

• R² - only for curve fitting model such as regression
  - In general, the lower the error measure (BIAS, MAD, MSE) or the higher the R², the better the forecasting model

Causal Forecasting

• Find a straight line that fits the data best.
  \[ Y = Intercept + slope * X (= a + bX) \]
  \[ Slope = \text{change in } Y / \text{change in } X \]
Causal Forecasting Models

- Curve Fitting: Simple Linear Regression
  - One Independent Variable \( (X) \) is used to predict one Dependent Variable \( (Y) \): \( Y = a + bX \)
  - Given \( n \) observations \( (X_i, Y_i) \), we can fit a line to the overall pattern of these data points. The Least Squares Method in statistics can give us the best \( a \) and \( b \) in the sense of minimizing \( \sum(Y_i - a - bX_i)^2 \):

\[
\begin{align*}
b &= \frac{\left( \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} \right) \left( \sum X_i^2 - \frac{(\sum X_i)^2}{n} \right)}{\left( \sum X_i^2 - \frac{(\sum X_i)^2}{n} \right)} \\
a &= \frac{\sum Y_i - b \sum X_i}{n}
\end{align*}
\]

Regression Example:

An oil company expansion

Consider an oil company that is planning to expand its network of modern self-service gasoline stations. The company plans to use traffic flow (measured in the average number of cars per hour) to forecast sales (measured in average dollar sales per hour). The firm has had five stations in operation for more than a year and has used historical data to calculate the following averages:

<table>
<thead>
<tr>
<th>Station</th>
<th>Cars per Hour</th>
<th>Sales per Hour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>150</td>
<td>220.00</td>
</tr>
<tr>
<td>2</td>
<td>55</td>
<td>75.00</td>
</tr>
<tr>
<td>3</td>
<td>220</td>
<td>250.00</td>
</tr>
<tr>
<td>4</td>
<td>130</td>
<td>145.00</td>
</tr>
<tr>
<td>5</td>
<td>95</td>
<td>200.00</td>
</tr>
</tbody>
</table>

Use Excel’s Data Analysis Tools | Regression
A linear trend is fit to the data:

Using Excel (2003 or older), click on Tools – Data Analysis …

In the resulting dialog, choose Regression.

Using Excel (2007), click on Data tab and then in the Analysis group, click Data Analysis Tools

In the Regression dialog, enter the Y-range and X-range.

Choose to place the output in a new worksheet called Results

Select Residual Plots and Normal Probability Plots to be created along with the output.
Click OK to produce the following results:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SUMMARY OUTPUT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Regression Statistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Multiple R</td>
<td>0.999937</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>R Square</td>
<td>0.693851</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Adjusted R Square</td>
<td>0.991930</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Standard Error</td>
<td>0.200766</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Observations</td>
<td>4.64.1704</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note that \( a \) (Intercept) and \( b \) (X Variable 1) are reported as 57.104 and 0.92997, respectively.

Forecasting Models

Now the question: Should we build a station at Buffalo Grove where traffic is 183 cars/hour?

The best guess at what the corresponding sales volume would be is found by placing this X value into the regression equation:

\[
\hat{Y} = a + b \times X
\]

Sales/hour = 57.104 + 0.92997 \times (183) = $227.29

Another value of interest in the Summary report is the t-statistic for the X variable and its associated values.

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Intercept</td>
<td>57.10423</td>
<td>0.38799983</td>
<td>1.13229</td>
<td>0.33947</td>
<td>-103.25311</td>
</tr>
<tr>
<td>18</td>
<td>X Variable 1</td>
<td>0.929976</td>
<td>0.259556467</td>
<td>2.60010</td>
<td>0.07801</td>
<td>-0.20475</td>
</tr>
</tbody>
</table>

The t-statistic is 2.61 and the P-value is 0.0798.

A P-value less than 0.05 represents that we have at least 95% confidence that the slope parameter \( (b) \) is statistically significantly than 0 (zero).

A slope of 0 results in a flat trend line and indicates no relationship between Y and X.

The 95% confidence limit for \( b \) is \([-0.205; 2.064]\).

Thus, we can’t exclude the possibility that the true value of \( b \) might be 0.
Time Series Model Building

- Historical data collection
- Data plotting (time series plot)
- Forecasting model building
- Evaluation and selection of model
- Forecasting with the final selected model

Components of A Time Series

- Trend: long term overall up or down movement
- Seasonality: periodic pattern repeating every year
- Cycles: up & down movement repeating over long time frame
- Random Variations: random movements follow no pattern
Time Series Forecasting Models

Look at the data (Time Series Plot)  
Forecast using one or more techniques  
Evaluate the technique and pick the best one.

<table>
<thead>
<tr>
<th>Observations from the time series plot</th>
<th>Techniques to try</th>
<th>Ways to evaluate</th>
</tr>
</thead>
</table>
| Data is reasonably stationary (no trend or seasonality) | Heuristics - Averaging methods  
- Naïve  
- Moving Averages  
- Simple Exponential Smoothing | • MAD  
• MAPE  
• MSE  
• BIAS |
| Data shows a consistent trend | Regression  
- Linear  
- Non-linear Regressions (not covered in this course) | • MAD  
• MAPE  
• MSE  
• BIAS  
• R-Squared |
| Data shows both a trend and a seasonal pattern | Classical decomposition  
- Find Seasonal Index  
- Use regression analyses to find the trend component | • MAD  
• MAPE  
• MSE  
• BIAS  
• R-Squared |

Trend Model

Curve fitting method used for time series data (also called time series regression model) Useful when the time series has a clear trend Can not capture seasonal patterns

Linear Trend Model: \( Y_t = a + bt \)
- \( t \) is time index for each period, \( t = 1, 2, 3, \ldots \)

Stationary Models

- Naïve
  - \textit{I sold 10 units yesterday, so I think I will sell 10 units today.}
- n-period moving average
  - \textit{For the past n days, I sold 12 units on average. Therefore, I think I will sell 12 units today.}
- Exponential smoothing
  - \textit{I predicted to sell 10 units at the beginning of yesterday; At the end of yesterday, I found out I actually sold 8 units. So, I will adjust the forecast of 10 (yesterday’s forecast) by adding adjusted error (\( \alpha \) \* error). This will compensate over- (or under-) forecast of yesterday.}

Naïve Model

The simplest time series forecasting model

Idea: “what happened last time (last year, last month, yesterday) will happen again this time”

Naïve Model:
- Algebraic: \( F_t = Y_{t-1} \)
  - \( Y_{t-1} \): actual value in period \( t-1 \)
  - \( F_t \): forecast for period \( t \)
- Spreadsheet: in B3: “= A2”; Copy down
Moving Average Model

Simple n-Period Moving Average

\[ F_t = \frac{\text{Sum of actual values in previous } n \text{ periods}}{n} \]

\[ = \frac{Y_{t-1} + Y_{t-2} + \cdots + Y_{t-n}}{n} \]

Issues of MA Model

- Naïve model is a special case of MA with \( n = 1 \)
- Idea is to reduce random variation or smooth data
- All previous \( n \) observation are treated equally (equal weights)
- Suitable for relatively stable time series with no trend or seasonal pattern

Forecasting Models 29

MA Example

<table>
<thead>
<tr>
<th>Month</th>
<th>Demand</th>
<th>( n=3 )</th>
<th>( n=5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>120</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Feb</td>
<td>90</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mar</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Apr</td>
<td>75</td>
<td>103.3</td>
<td>-</td>
</tr>
<tr>
<td>May</td>
<td>110</td>
<td>88.3</td>
<td>-</td>
</tr>
<tr>
<td>June</td>
<td>50</td>
<td>95.0</td>
<td>99.0</td>
</tr>
<tr>
<td>July</td>
<td>75</td>
<td>78.3</td>
<td>85.0</td>
</tr>
<tr>
<td>Aug</td>
<td>130</td>
<td>78.3</td>
<td>82.0</td>
</tr>
<tr>
<td>Sept</td>
<td>110</td>
<td>85.0</td>
<td>88.0</td>
</tr>
<tr>
<td>Oct</td>
<td>90</td>
<td>105.0</td>
<td>95.0</td>
</tr>
<tr>
<td>Nov</td>
<td>?</td>
<td>110.0</td>
<td>91.0</td>
</tr>
</tbody>
</table>

Forecasting Models 30

Smoothing Effect of MA Model

Longer-period moving averages (larger \( n \)) react to actual changes more slowly or will get less impact from a new data point

Forecasting Models 31

Moving Average Model

Weighted n-Period Moving Average

\[ F_t = w_1 Y_{t-1} + w_2 Y_{t-2} + \cdots + w_n Y_{t-n} \]

- Typically weights are decreasing: \( w_1 > w_2 > \cdots > w_n \)
- Sum of the weights = \( \sum w_i = 1 \)
- Flexible weights reflect relative importance of each previous observation in forecasting
- Optimal weights can be found via Solver

Forecasting Models 32
Weighted MA: An Illustration

<table>
<thead>
<tr>
<th>Month</th>
<th>Weight</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>August</td>
<td>17%</td>
<td>130</td>
</tr>
<tr>
<td>September</td>
<td>33%</td>
<td>110</td>
</tr>
<tr>
<td>October</td>
<td>50%</td>
<td>90</td>
</tr>
</tbody>
</table>

November forecast:

\[ F_{Nov} = (0.50)(90) + (0.33)(110) + (0.17)(130) \]

\[ = 103.4 \]

Simple Exponential Smoothing

A special type of weighted moving average

- Include all past observations
- Use a unique set of weights that weight recent observations much more heavily than very old observations:

\[ 0 < \alpha < 1 \]

\[ \text{Decreasing weights given to older observations} \]

Simple ES: The Model

\[ F_t = \alpha Y_{t-1} + \alpha (1 - \alpha) Y_{t-2} + \alpha (1 - \alpha)^2 Y_{t-3} + \cdots \]

\[ F_t = \alpha Y_{t-1} + (1 - \alpha) [\alpha Y_{t-2} + \alpha (1 - \alpha) Y_{t-3} + \cdots ] \]

\[ F_t = \alpha Y_{t-1} + (1 - \alpha) F_{t-1} \]

\[ = F_{t-1} + \alpha (Y_{t-1} - F_{t-1}) \]

- \( \alpha \): Smoothing constant
- \( F_t \): Forecast for period \( t \)
- \( F_{t-1} \): Last period forecast
- \( Y_{t-1} \): Last period actual value

ES Example

<table>
<thead>
<tr>
<th>Period</th>
<th>Month</th>
<th>Demand</th>
<th>Forecast ( \alpha = 0.1 )</th>
<th>Forecast ( \alpha = 0.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan</td>
<td>37</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>2</td>
<td>Feb</td>
<td>40</td>
<td>37.00</td>
<td>37.00</td>
</tr>
<tr>
<td>3</td>
<td>Mar</td>
<td>41</td>
<td>37.30</td>
<td>38.50</td>
</tr>
<tr>
<td>4</td>
<td>Apr</td>
<td>37</td>
<td>37.67</td>
<td>39.75</td>
</tr>
<tr>
<td>5</td>
<td>May</td>
<td>45</td>
<td>37.60</td>
<td>38.37</td>
</tr>
<tr>
<td>6</td>
<td>Jun</td>
<td>50</td>
<td>38.34</td>
<td>41.68</td>
</tr>
<tr>
<td>7</td>
<td>Jul</td>
<td>43</td>
<td>39.51</td>
<td>45.84</td>
</tr>
<tr>
<td>8</td>
<td>Aug</td>
<td>47</td>
<td>39.86</td>
<td>44.42</td>
</tr>
<tr>
<td>9</td>
<td>Sep</td>
<td>56</td>
<td>40.57</td>
<td>45.71</td>
</tr>
<tr>
<td>10</td>
<td>Oct</td>
<td>52</td>
<td>42.11</td>
<td>50.85</td>
</tr>
<tr>
<td>11</td>
<td>Nov</td>
<td>55</td>
<td>43.10</td>
<td>51.42</td>
</tr>
<tr>
<td>12</td>
<td>Dec</td>
<td>54</td>
<td>44.29</td>
<td>53.21</td>
</tr>
<tr>
<td>13</td>
<td>Jan</td>
<td>?</td>
<td>45.26</td>
<td>53.61</td>
</tr>
</tbody>
</table>

Assume \( F_1 = Y_1 \)

\[ = 0.1(37) + 0.9(37) \]

\[ = 0.1(40) + 0.9(37.00) \]
The Smoothening Effect of $\alpha$

![Graph showing the smoothing effect of different $\alpha$ values]

Larger $\alpha$ reacts to actual changes more quickly (responsive) while smaller $\alpha$ responds more slowly to actual changes.

Properties of Simple ES

- Widely used and successful model
- Requires very little data
- Larger $\alpha$, more responsive forecast; Smaller $\alpha$, smoother forecast (See also Table 13.2)
- “Best” $\alpha$ in terms of minimizing MSE or MAD can be found by Solver

Time Series Decomposition Model

Basic Idea: a time series is composed of several basic components: Trend, Seasonality, Cycle, and Random Error

The multiplicative decomposition model:

$$Y_t = \text{Trend}_t \times \text{Cycle}_t \times \text{Seasonality}_t \pm \text{Error}_t$$

- These components contribute to time series value in a multiplicative way

Time Series Decomposition

The basic model is:

$$Y = \text{Trend} \times \text{Cyclical} \times \text{Seasonal} \pm \text{Error}$$

Since we cannot easily extract or predict cycles, we will assume that the trend component will capture cycles during the forecast period.

Since we have to live with error (cannot predict it), our model is simplified to:

- $Y = \text{Trend} \times \text{Seasonal}$
I. Estimate *Seasonal Index* (Simplified method)

1. Calculate overall average demand using all data points
2. Divide each demand by overall demand average
   - Each resulting number is called a **raw index**
   - The number of raw indices for the same month or quarter is equal to the number of years
3. Calculate Seasonal Index (SI) by averaging all raw indices for the same month or quarter

### Example: average of year 1
- January ratio: \((0.851 + 1.064)/2 = 0.957\)
- January ratio: \((0.851 + 1.064)/2 = 0.957\)

II. Estimate *Trend* Component

Step 1: Remove seasonal effect
- Deseasonalized data\(_t\) = \(Y_t / SI_t\)

Step 2: Fit a trend line to deseasonalized data using least squares method

Step 3: Calculate the trend value for each period

- **Note:** If the deseasonalized data look stable (no apparent trend), simple exponential smoothing may be used in Steps 2 and 3 to calculate the forecast (rather than trend) for each period.

III. Forecast

Combine seasonal and trend components

- \(F_t = \text{Trend Value}_t \times \text{Seasonal Index}_t\)

- This final step is also called reseasonalizing
- Trend Value\(_t\) is the trend estimate for the period \(t\), based on the trend model fitted to the deseasonalized data