Demand Forecasting

Lectures 2 & 3
ESD.260 Fall 2003

Caplice

Agenda

- The Problem and Background
- Four Fundamental Approaches
- Time Series
  - General Concepts
  - Evaluating Forecasts – How 'good' is it?
  - Forecasting Methods (Stationary)
    - Cumulative Mean
    - Naïve Forecast
    - Moving Average
    - Exponential Smoothing
  - Forecasting Methods (Trends & Seasonality)
    - OLS Regression
    - Holt's Method
    - Exponential Method for Seasonal Data
    - Winter's Model
- Other Models

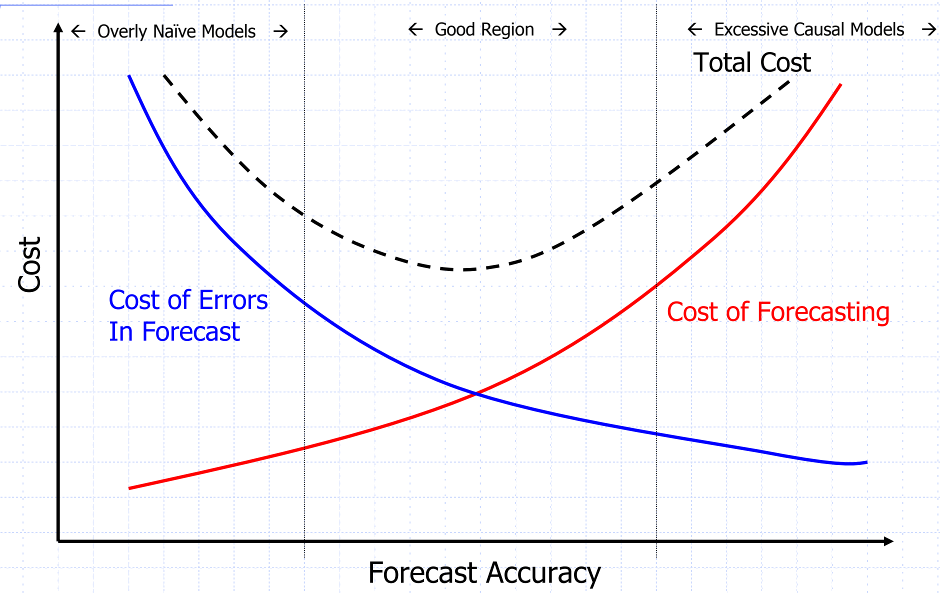
Demand Forecasting

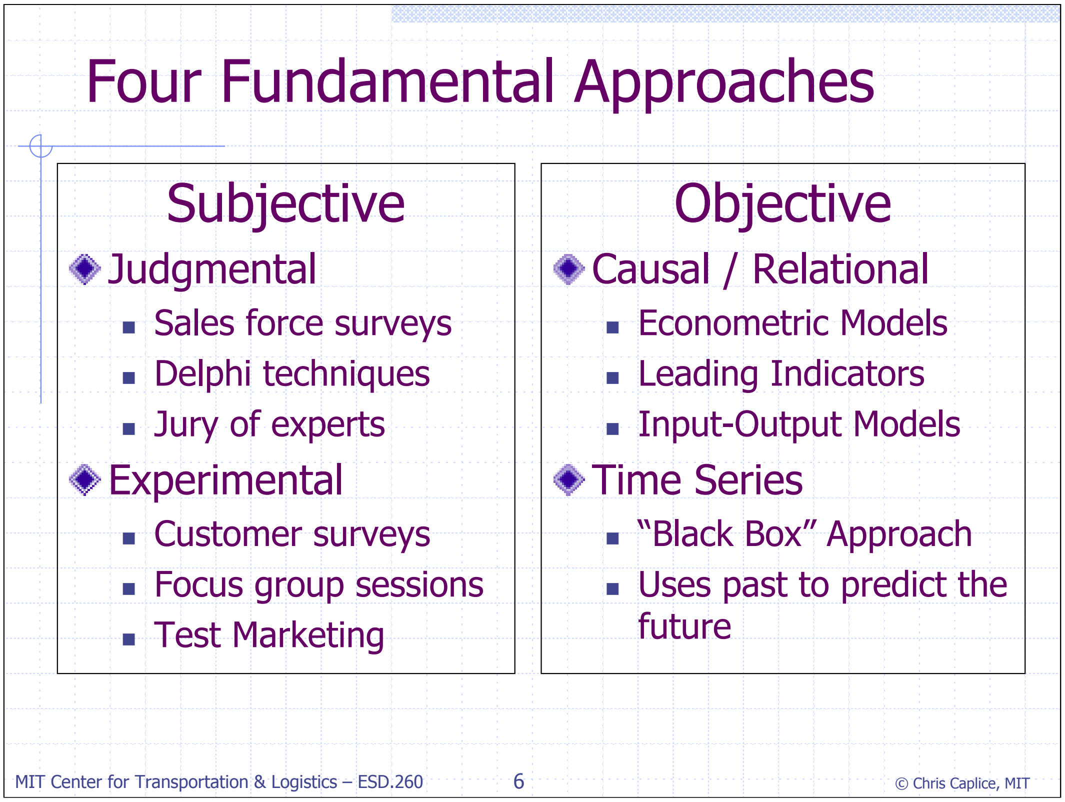
- The problem:
  - Generate the large number of short-term, SKU level, locally dis-aggregated demand forecasts required for production, logistics, and sales to operate successfully.
  - Focus on:
    - Forecasting product demand
    - Mature products (not new product releases)
    - Short time horiDon (weeks, months, quarters, year)
    - Use of models to assist in the forecast
    - Cases where demand of items is independent

Demand Forecasting – Punchline(s)

- Forecasting is difficult – especially for the future
- Forecasts are always wrong
- The less aggregated, the lower the accuracy
- The longer the time horizon, the lower the accuracy
- The past is usually a pretty good place to start
- Everything exhibits seasonality of some sort
- A good forecast is not just a number – it should include a range, description of distribution, etc.
- Any analytical method should be supplemented by external information
- A forecast for one function in a company might not be useful to another function (Sales to Mkt to Mfg to Trans)

Cost of Forecasting vs Inaccuracy





S – used by sales/marketing

O – used by prod and inv

All are a search for pattern

Judge:

Exp: new products then extrapolate

Causal: sports jerseys, umbrellas

Time Series: different – is not looking for cause or judgement, only repeating patterns

Time Series Concepts

1. Time Series – Regular & recurring basis to forecast
2. Stationarity – Values hover around a mean
3. Trend – Persistent movement in one direction
4. Seasonality – Movement periodic to calendar
5. Cycle – Periodic movement not tied to calendar
6. Pattern + Noise – Predictable and random components of a Time Series forecast
7. Generating Process –Equation that creates TS
8. Accuracy and Bias – Closeness to actual vs Persistent tendency to over or under predict
9. Fit versus Forecast – Tradeoff between accuracy to past forecast to usefulness of predictability
10. Forecast Optimality – Error is equal to the random noise

Time Series - A set of numbers which are observed on a regular, recurring basis. Historical observations are known; future values must be forecasted.

Stationarity - Values of a series hover or cluster around a fixed mean, or level.

Trend - Values of a series show persistent movement in one direction, either up or down. Trend may be linear or non-linear.

Seasonality - Values of a series move up or down on a periodic basis which can be related to the calendar.

Cycle - Values of a series move through long term upward and downward swings which are not related to the calendar.

Pattern + Noise - A Time Series can be thought of as two components: Pattern, which can be used to forecast, and Noise, which is purely random and cannot be forecasted.

Generating Process - The "equation" which actually creates the time series observations. In most real situations, the generating process is unknown and must be inferred.

Accuracy and Bias - Accuracy is a measure of how closely forecasts align with observations of the series. Bias is a persistent tendency of a forecast to over-predict or under-predict. Bias is therefore a kind of pattern which suggests that the procedure being used is inappropriate.

Fit versus Forecast - A forecast model which has been "tuned" to fit a historical data set very well will not necessarily forecast future observations more accurately. A more complicated model can always be devised which will fit the old data well -- but which will probably work poorly with new observations.

Forecast Optimality - A forecast is optimal if all the actual pattern in the process has been discovered. As a result, all remaining forecast error is attributable to "unforecastable" noise. In more technical terms, the forecast is optimal if the mean squared forecast error equals the variance of the noise term in the long run. Note that an optimal forecast is not necessarily a "perfect" forecast. Some forecast error is expected to occur.

Some people purposely over-bias – bad idea – do not use forecasting to do inventory planning

Forecast Evaluation

- Measure of Accuracy / Error Estimates
- Notation:

  D_t = Demand observed in the t^{th} period

  F_t = Forecasted demand in the t^{th} period at the end of the $(t-1)^{th}$ period

  n = Number of periods

Capture the model and ignore the noise

Accuracy and Bias Measures

1. Forecast Error: $e_t = D_t - F_t$

2. Mean Deviation: $$MD = \frac{\sum_{t=1}^{n} e_t}{n}$$

3. Mean Absolute Deviation $$MAD = \frac{\sum_{t=1}^{n} |e_t|}{n}$$

4. Mean Squared Error: $$MSE = \frac{\sum_{t=1}^{n} e_t^2}{n}$$

5. Root Mean Squared Error: $$RMSE = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}}$$

6. Mean Percent Error: $$MPE = \frac{\sum_{t=1}^{n} \frac{e_t}{D_t}}{n}$$

7. Mean Absolute Percent Error: $$MAPE = \frac{\sum_{t=1}^{n} \frac{|e_t|}{D_t}}{n}$$

MIT Center for Transportation & Logistics – ESD.260 9

MD – cancels out the over and under – good measure of bias not accuracy

MAD – fixes the cancelling out, but statistical properties are not suited to probability based dss

MSE – fixes cancelling out, equivalent to variance of forecast errors, HEAVILY USED statistically appropriate measure of forecast errors

RMSE – easier to interpret (proportionate in large data sets to MAD) MAD/RMSE = SQRT(2/pi) for e~N

Relative metrics are weighted by the actual demand

MPE – shows relative bias of forecasts

MAPE – shows relative accuracy

Optimal is when the MSE of forecasts -> Var(e) – thus the forecsts explain all but the noise.

What is good in practice (hard to say) MAPE 10% to 15% is excellent, MAPE 20%-30% is average CLASS?

The Cumulative Mean

Generating Process:

$$D_t = L + n_t$$

where: $n_t \sim iid\ (\mu = 0\ ,\ \sigma^2 = V[n])$

Forecasting Model:

$$F_{t+1} = (D_1 + D_2 + D_3 + ... + D_t) / t$$

MIT Center for Transportation & Logistics – ESD.260 10 © Chris Caplice, MIT

Stationary model – mean does not change – pattern is a constant

Not used in practice – is anything constant?

Thought though is to use as large a sample siDe as possible to

The Naïve Forecast

Generating Process:

$$D_t = D_{t-1} + n_t$$

where: $n_t \sim \text{iid}\ (\mu = 0\ ,\ \sigma^2 = V[n])$

Forecasting Model:

$$F_{t+1} = D_t$$

The Moving Average

Generating Process:

$$D_t = L + n_t \quad ; \; t < t_s$$

$$D_t = L + S + n_t \quad ; \; t \geq t_s$$

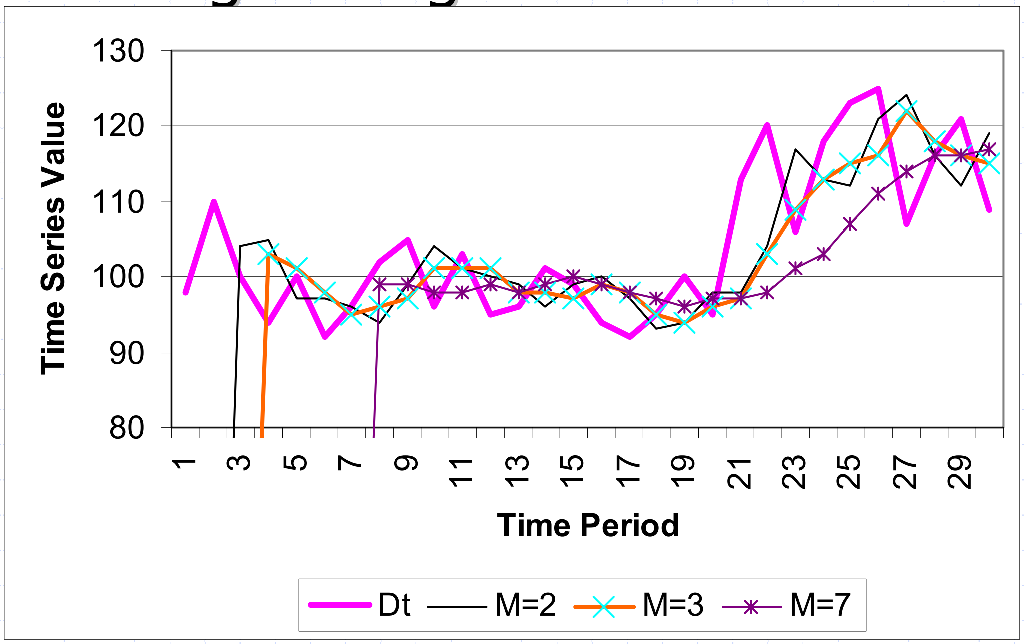
where: $n_t \sim \text{iid}\,(\mu = 0 \,,\, \sigma^2 = V[n])$

Forecasting Model:

$$F_{t+1} = (D_t + D_{t-1} + \ldots + D_{t-M+1}) / M$$

where M is a parameter

Moving Average Forecasts



Exponential Smoothing

$$F_{t+1} = \alpha D_t + (1-\alpha) F_t$$

Where: $0 < \alpha < 1$

An Equivalent Form:

$$F_{t+1} = F_t + \alpha e_t$$

Another way to think about it...

$F_{t+1} = \alpha D_t + (1-\alpha)F_t$ *but recall that* $F_t = \alpha D_{t-1} + (1-\alpha)F_{t-1}$

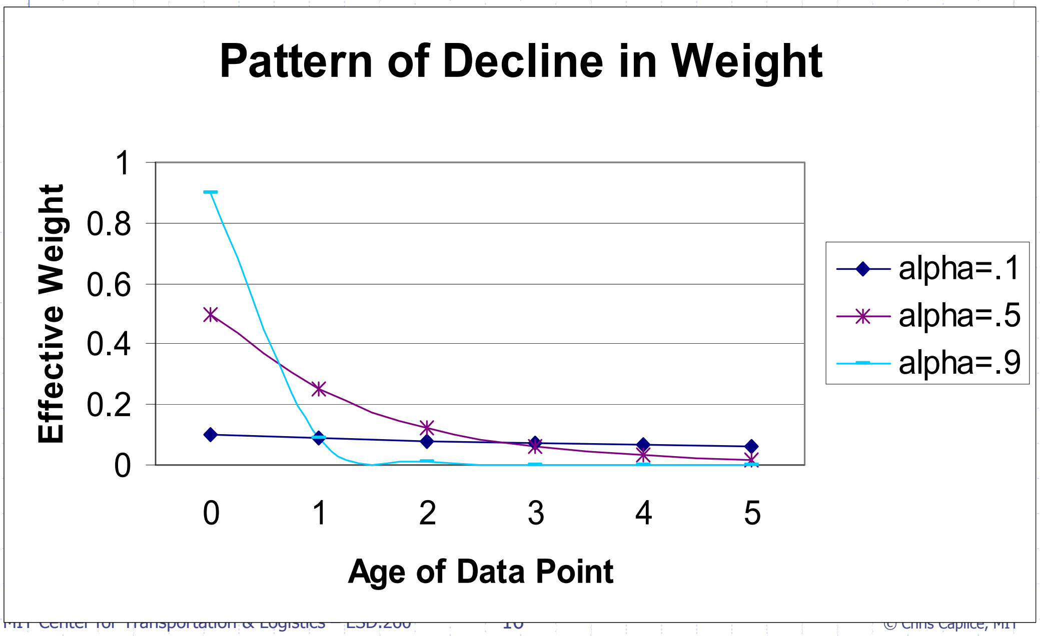
$F_{t+1} = \alpha D_t + (1-\alpha)(\alpha D_{t-1} + (1-\alpha)F_{t-1})$

$F_{t+1} = \alpha D_t + \alpha(1-\alpha)D_{t-1} + (1-\alpha)^2F_{t-1}$

$F_{t+1} = \alpha D_t + \alpha(1-\alpha)D_{t-1} + \alpha(1-\alpha)^2D_{t-2} + (1-\alpha)^3F_{t-3}$

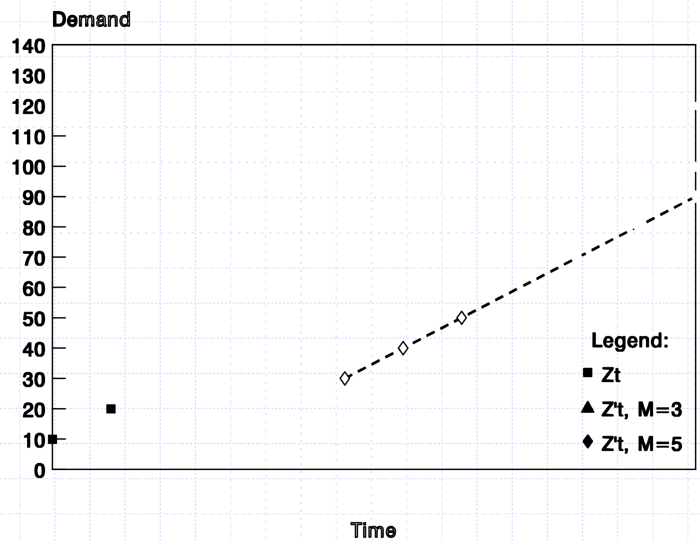
$F_{t+1} = \alpha(1-\alpha)^0D_t + \alpha(1-\alpha)^1D_{t-1} + \alpha(1-\alpha)^2D_{t-2} + \alpha(1-\alpha)^3D_{t-3}...$

Exponential Smoothing



Forecasting Trended Data

Why not to use stationary models on trended



Holt's Model for Trended Data

Forecasting Model:

$$F_{t+1} = L_{t+1} + T_{t+1}$$

Where: $L_{t+1} = \alpha D_t + (1-\alpha)(L_t + T_t)$

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

Exponential Smoothing for Seasonal Data

Forecasting Model: $F_{t+1} = L_{t+1} S_{t+1-m}$

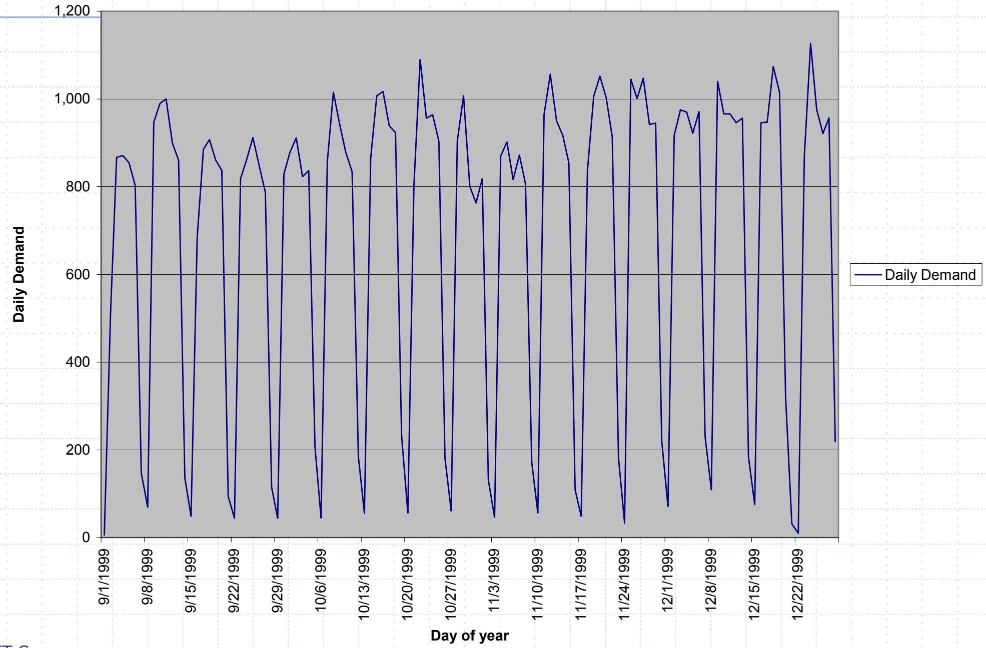
Where : $L_{t+1} = \alpha(D_t/S_t) + (1-\alpha) L_t$

And : $S_{t+1} = \gamma(D_{t+1}/L_{t+1}) + (1-\gamma) S_{t+1-m}$,

An Example where m=4, t=12:

$$F_{13} = L_{13} S_9$$
$$L_{13} = \alpha(D_{12}/S_{12}) + (1-\alpha) L_{12}$$
$$S_{13} = \gamma (D_{13}/L_{13}) + (1-\gamma) S_9$$

Seasonal Pattern



MIT Ce MIT

Winter's Model for Trended/Seasonal Data

$$F_{t+1} = (L_{t+1} + T_{t+1})\ S_{t+1-m}$$

$$L_{t+1} = \alpha(D_t/S_t) + (1-\alpha)(L_t + T_t)$$

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

$$S_{t+1} = \gamma(D_{t+1}/L_{t+1}) + (1-\gamma)\ S_{t+1-m}$$

Causal Approaches

- Econometric Models
  - There are underlying reasons for demand
  - Examples:
- Methods:
  - Continuous Variable Models
    - Ordinary Least Squares (OLS) Regression
    - Single & Multiple Variable
  - Discrete Choice Methods
    - Logit & Probit Models
    - Predicting demand for making a choice

Other Issues in Forecasting

- Data Issues: Demand Data ≠ Sales History
- Model Monitoring
- Box-Jenkins ARIMA Models, etc.
- Focus Forecasting
- Use of Simulation
- Forecasting Low Demand Items
- Collaborative Planning, Forecasting, and Replenishment (CPFR)
- Forecasting and Inventory Management