Forecasting Models – Chapter 2

IE 3265

R. Lindeke, Ph. D.

Introduction to Forecasting

- What is forecasting?
  - Primary Function is to Predict the Future using (time series related or other) data we have in hand
- Why are we interested?
  - Affects the decisions we make today
- Where is forecasting used in POM
  - forecast demand for products and services
  - forecast availability/need for manpower
  - forecast inventory and materiel needs daily

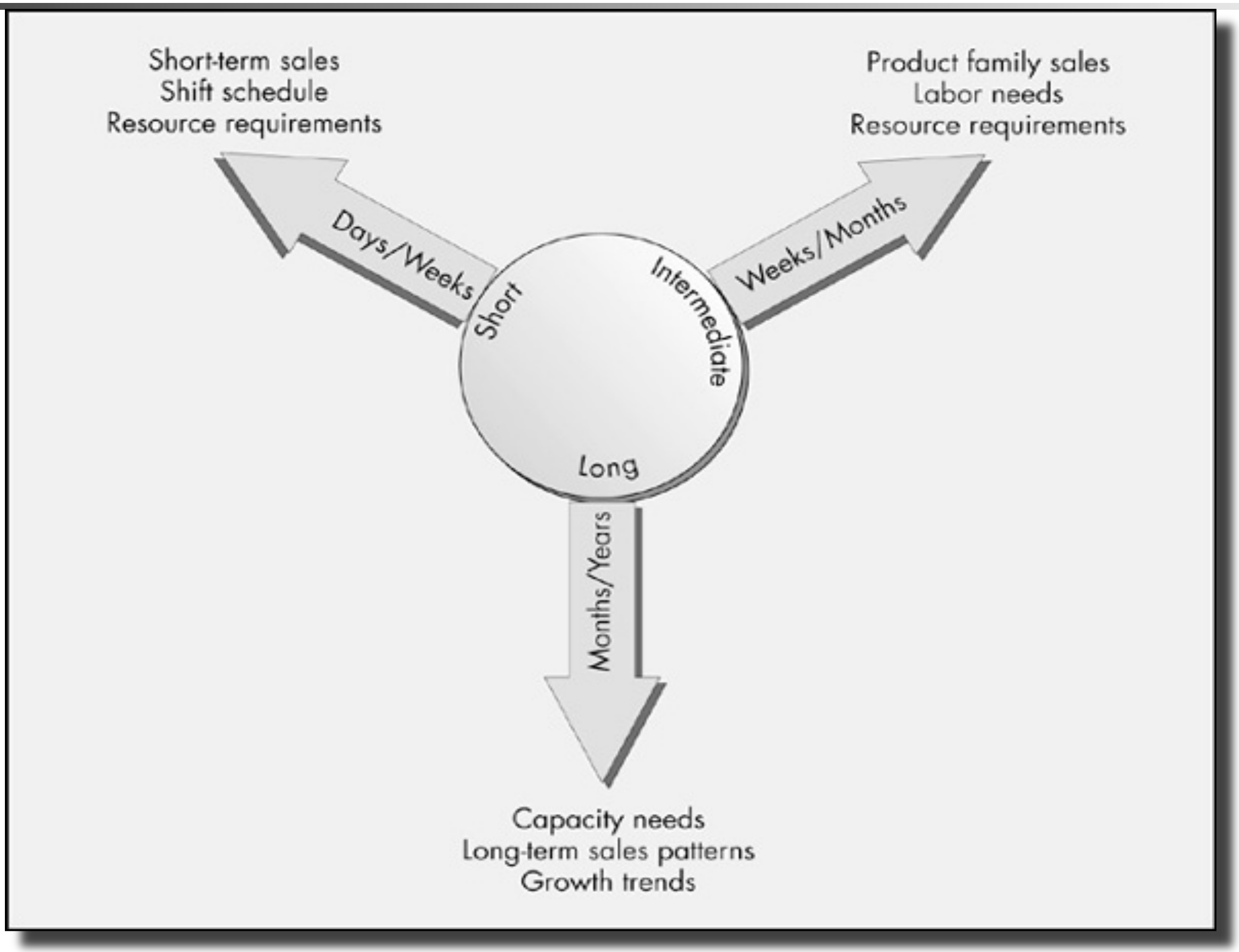
Characteristics of Forecasts

- They are usually wrong!
- A good forecast is more than a single number
  - Includes a mean value and standard deviation
  - Includes accuracy range (high and low)
- Aggregated forecasts are usually more accurate
- Accuracy erodes as we go further into the future.
- Forecasts *should not* be used to the exclusion of known information

What Makes a Good Forecast?

- It should be timely
- It should be as accurate as possible
- It should be reliable
- It should be in meaningful units
- It should be presented in writing
- The method should be easy to use and understand in most cases.

Forecast Horizons in Operation Planning – Figure 2.1



Subjective Forecasting Methods

- Sales Force Composites
  - Aggregation of sales personnel estimates
- Customer Surveys
- Jury of Executive Opinion
- The Delphi Method
  - Individual opinions are compiled and considered. These are anonymously shared among group. Then opinion request is repeated until an overall group consensus is (hopefully) reached.

Objective Forecasting Methods

- Two primary methods: *causal models* and *time series methods*

Causal Models

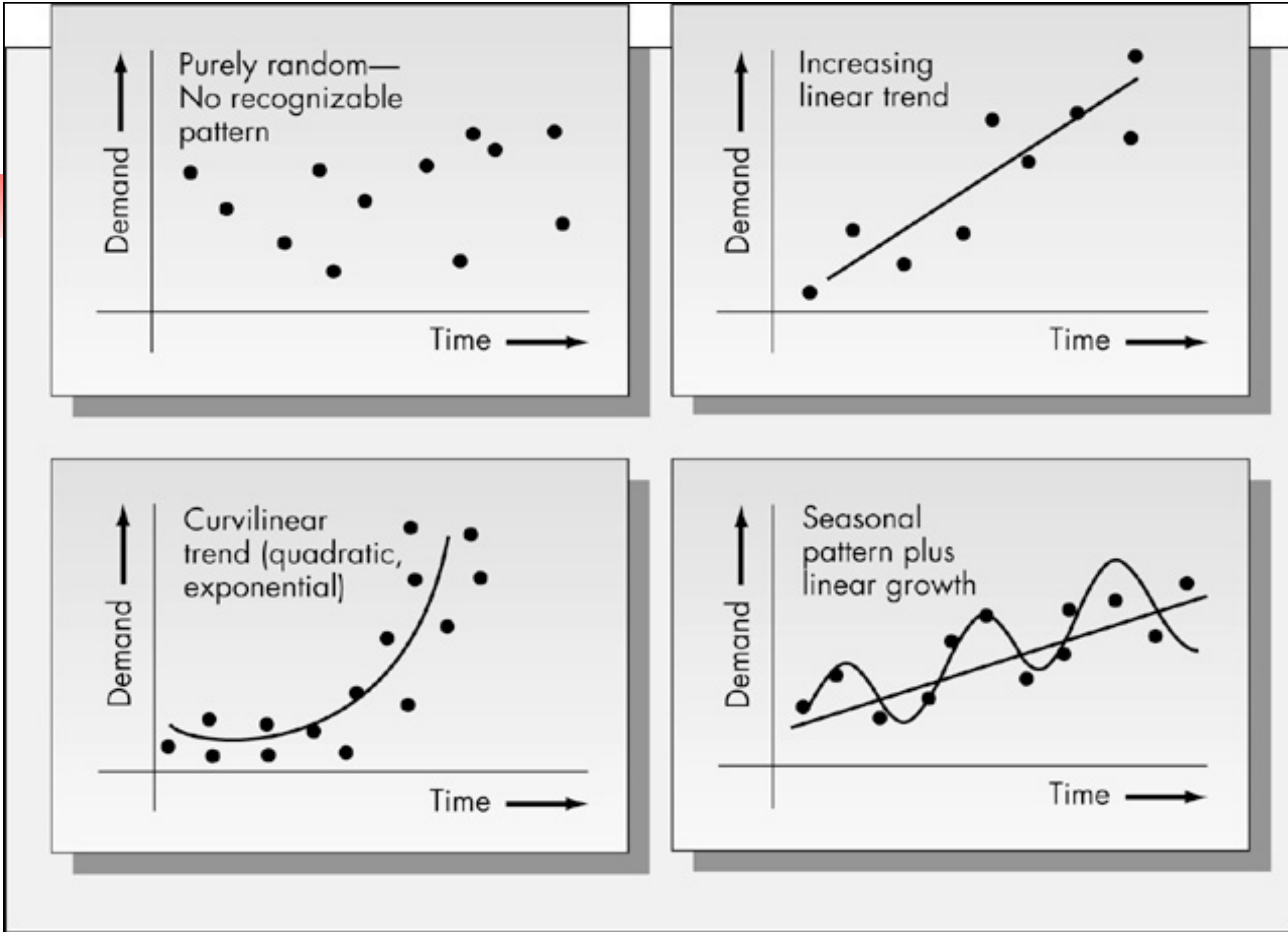
Let Y be the quantity to be forecasted and $(X_1, X_2, \ldots, X_n)$ are n variables that have predictive power for Y. A causal model is $Y = f(X_1, X_2, \ldots, X_n)$.

A typical relationship is a linear one:

$$Y = a_0 + a_1 X_1 + \ldots + a_n X_n$$

Time Series Methods

- A time series is just collection of past values of the variable being predicted. Also known as naïve methods. Goal is to isolate patterns in past data. (See Figures on following pages)
  - Trend
  - Seasonality
  - Cycles
  - Randomness


Purely random—
No recognizable
pattern
Demand
Time
Increasing
linear trend
Demand
Time
Curvilinear
trend (quadratic,
exponential)
Demand
Time
Seasonal
pattern plus
linear growth
Demand
Time

Notation Conventions

- Let $D_1, D_2, \ldots D_n, \ldots$ be the past values of the series to be predicted (demands?). If we are making a forecast during period t (for the future), assume we have observed D_t , D_{t-1} etc.
  - Let $F_{t, t + \tau}$ = forecast made in period t for the demand in period $t + \tau$ where $\tau = 1, 2, 3, \ldots$
  - Then $F_{t -1, t}$ is the forecast made in t-1 for t and
  - $F_{t, t+1}$ is the forecast made in t for t+1. (one step ahead) Use shorthand notation $F_t = F_{t - 1, t}$

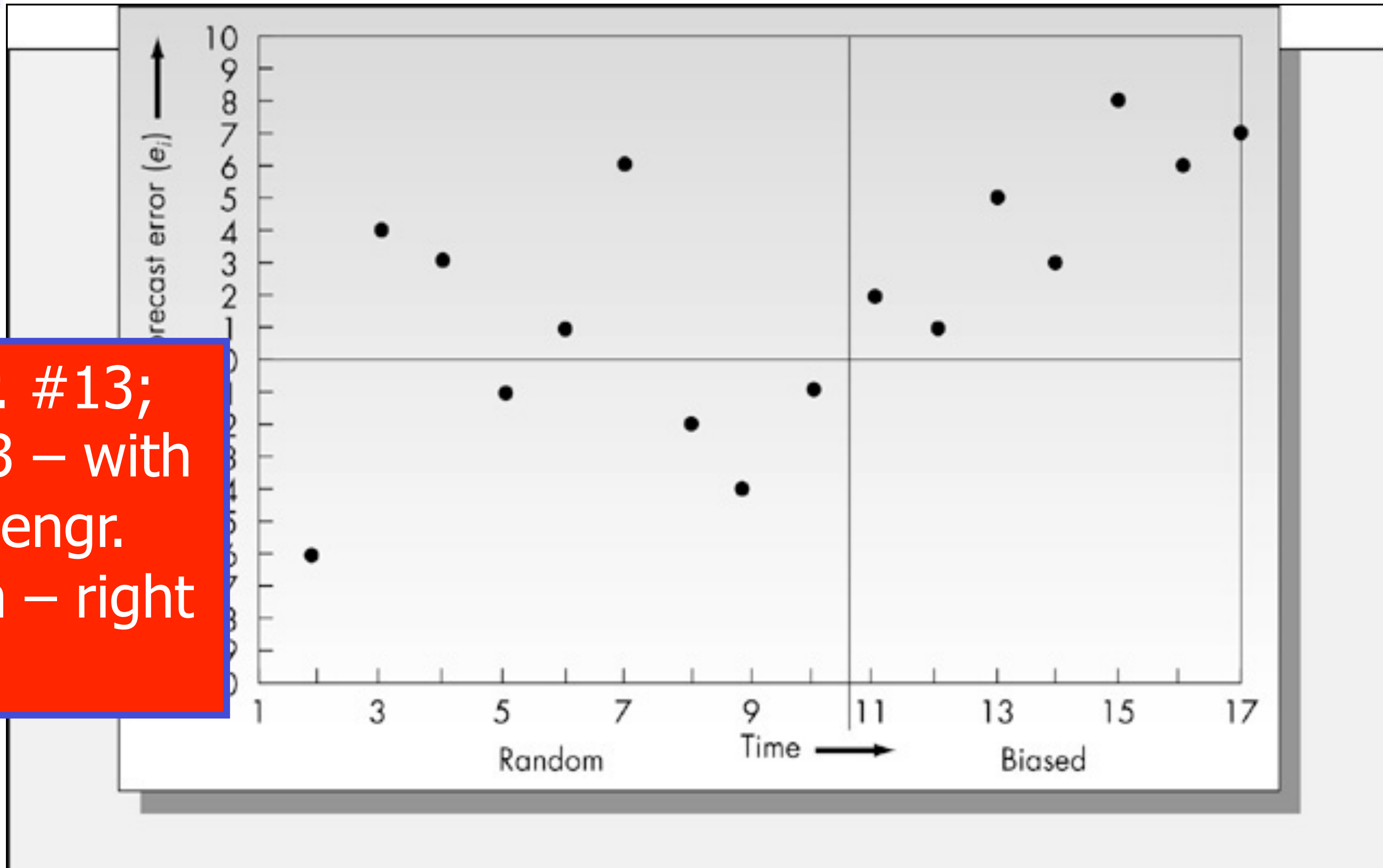
Evaluation of Forecasts

- The forecast error in period t, e_t, is the difference between the forecast for demand in period t and the actual value of demand in t.
- For a multiple step ahead forecast: $e_t = F_{t-\tau, t} - D_t$.
- For one step ahead forecast: $e_t = F_t - D_t$
- To evaluate Forecasting accuracy we develop a chart of Forecasting errors using:
  - $MAD = (1/n) \Sigma | e_i |$
  - $MSE = (1/n) \Sigma e_i^2$

Evaluation of Forecasts

- We would find that the value $MSE^{0.5} = 1.25MAD = \sigma_{error}$
- We can set up a "Control Chart" for tracking errors to study acceptability of our model of the forecast
- "Control" limits can be set at $\pm 3(\sigma_{error})$ or $\pm 3.75MAD$
- This chart can be used to study bias (trends that deviate in only one direction for some time indicating lags) or very large errors (or growing) indicating a need to review the models

Forecast Errors Over Time – Figure 2.3



Try Pr. #13; pg 63 – with your engr. Team – right now!

Forecasting for Stationary Series

- A stationary time series has the form:
  $D_t = \mu + \varepsilon_t$
  where μ is a constant and ε_t is a random variable with mean 0 and var σ^2
- Stationary series indicate stable processes without observable trends
- Two common methods for forecasting stationary series are moving averages and exponential smoothing.

Moving Averages

- In words: the arithmetic average of the n most recent observations. For a one-step-ahead forecast:

  $F_t = (1/N)\ (D_{t-1} + D_{t-2} + \ldots + D_{t-n})$

(Go to Example. Try 2.16 on page 66)

Moving Average -- example

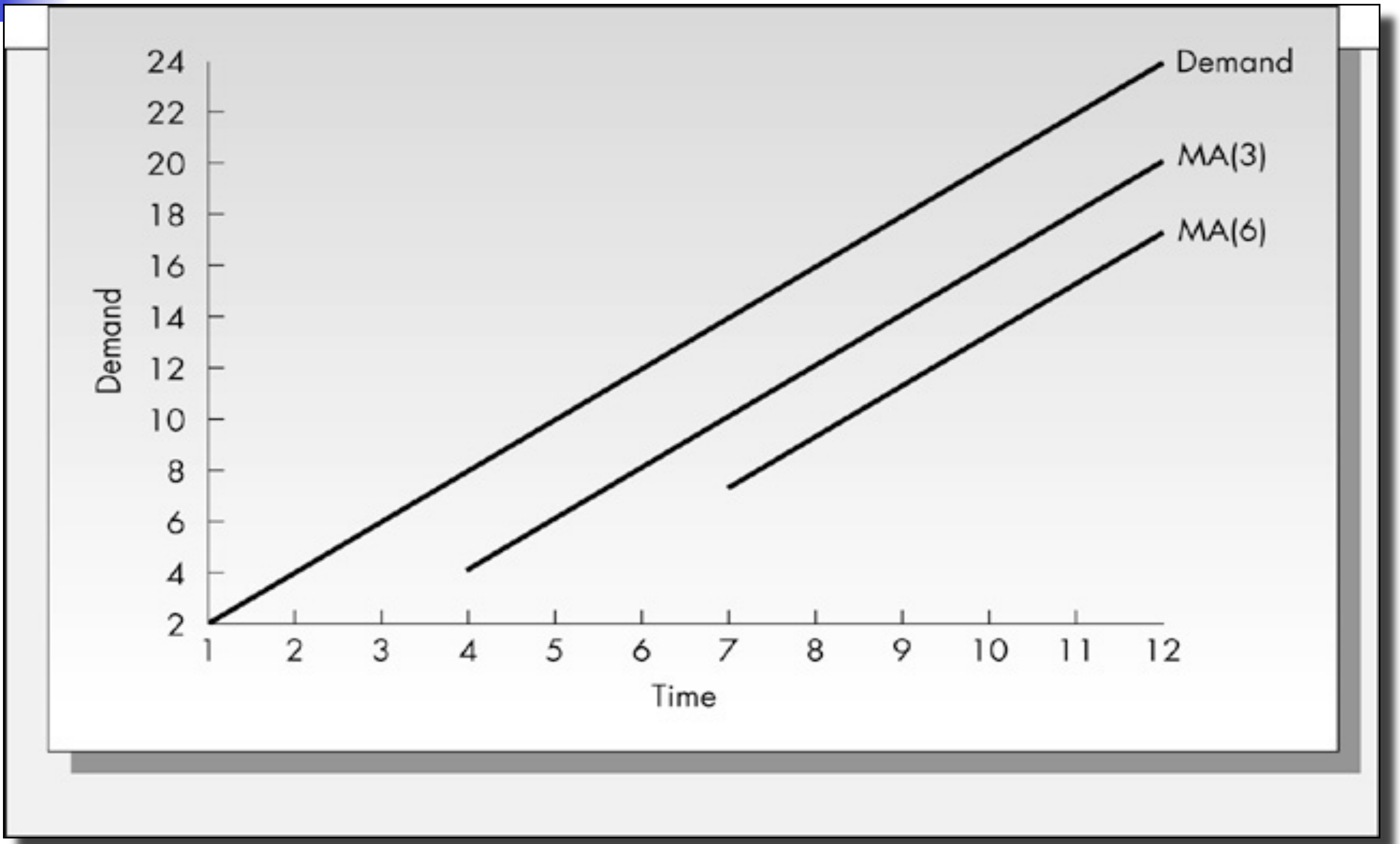
MONTH	Demand	Month	Demand
January	89	July	223
February	57	August	286
March	144	September	212
April	221	October	275
May	177	November	188
June	280	December	312

- 3 month MA: (oct+nov+dec)/3=258.33
- 6 month MA: (jul+aug+…+dec)/6=249.33
- 12 month MA: (Jan+feb+…+dec)/12=205.33

Summary of Moving Averages

- *Advantages of Moving Average Method*
  - Easily understood
  - Easily computed
  - Provides stable forecasts
- *Disadvantages of Moving Average Method*
  - Requires saving lots of past data points: at least the N periods used in the moving average computation
  - Lags behind a trend
  - Ignores complex relationships in data

Moving Average Lags a Trend – Figure 2.4



What about Weighted Moving Averages?

- This method looks at past data and tries to logically attach importance to certain data over other data
- Weighting factors must add to one
- Can weight recent higher than older or specific data above others
  - If forecasting staffing, we could use data from the last four weeks where Tuesdays are to be forecast.
  - Weighting on Tuesdays is: T_{-1} is .25; T_{-2} is .20; T_{-3} is .15; T_{-4} is .10 and Average of all other days is weighed .30.

Exponential Smoothing Method

- A type of weighted moving average that applies declining weights to past data.

1. New Forecast = α (most recent observation) + (1 - α) (last forecast)

   - OR -

2. New Forecast = last forecast - α (last forecast error)

where $0 < \alpha < 1$ and generally is small for stability of forecasts (around .1 to .2)

Exponential Smoothing (cont.)

In symbols:

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

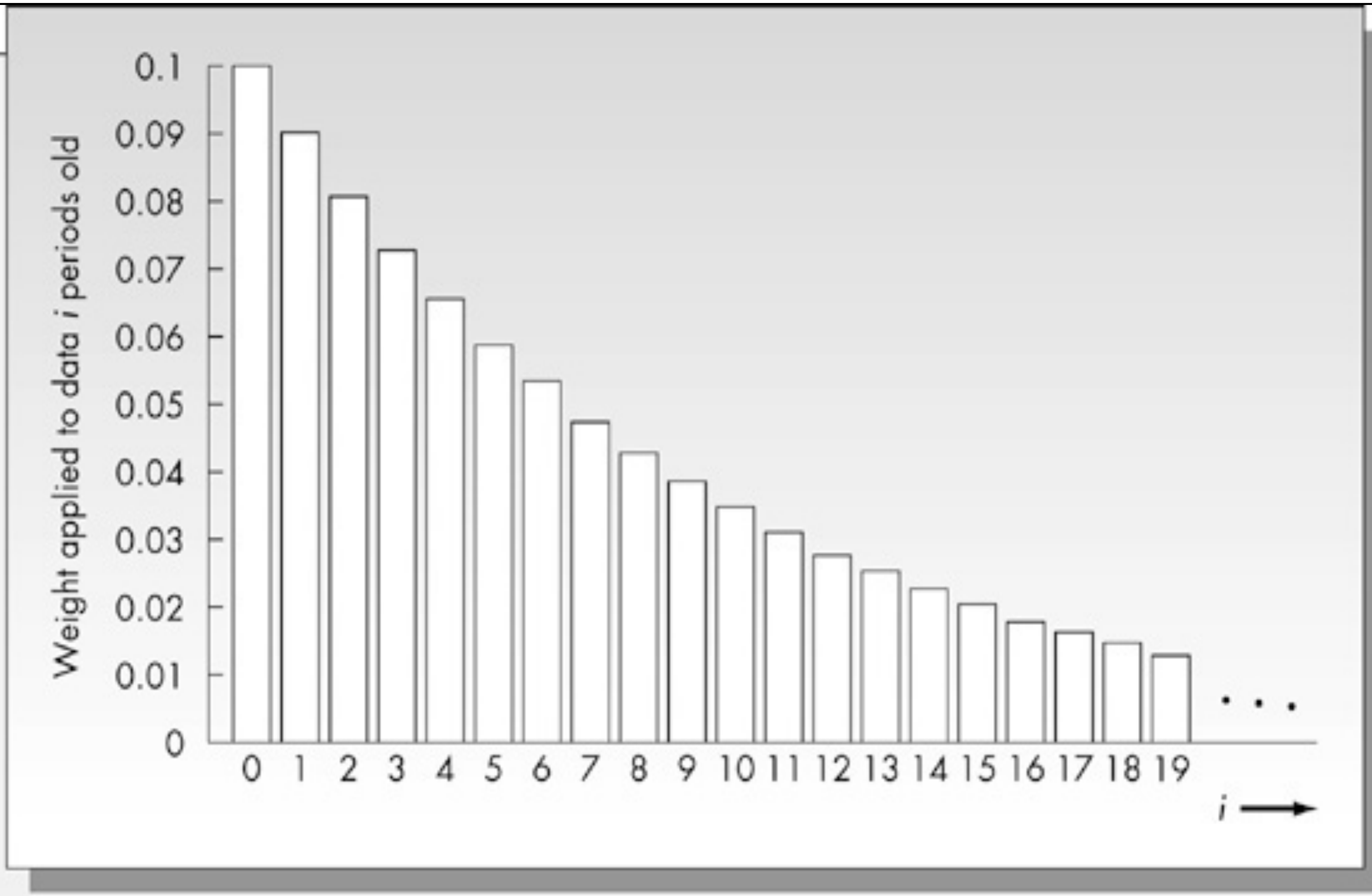
$$= \alpha D_t + (1 - \alpha)(\alpha D_{t-1} + (1 - \alpha) F_{t-1})$$

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

- Hence the method applies a set of exponentially declining weights to past data. It is easy to show that the *sum of the weights* is exactly one.

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

Weights in Exponential Smoothing:



Effect of α value on the Forecast

- Small values of α means that the forecasted value will be stable (show low variability
  - Low α increases the lag of the forecast to the actual data if a trend is present
- Large values of α mean that the forecast will more closely track the actual time series

Lets try one:

- #22a, pg 72: Given sales history of
  - Jan 23.3 Feb 72.3 Mar 30.3 Apr 15.5
  - And the January Forecast was: 25
  - Using α = .15
  - Forecast for Feb: $\alpha D_{jan} + (1-\alpha)F_{jan}$ = .15*23.3 + (.85)*25 = 24.745
  - Forecast for Mar: $\alpha D_{feb} + (1-\alpha)F_{feb}$ = .15*72.3 + (.85)*24.745 = 31.88
  - Apr: $\alpha D_{mar} + (1-\alpha)F_{mar}$ = .15*30.3 + .85*31.88 = 31.64
  - May: $\alpha D_{apr} + (1-\alpha)F_{apr}$ = .15*15.5 + .85*31.64 = 29.22

Comparison of MA and ES

- Similarities
  - Both methods are appropriate for stationary series
  - Both methods depend on a single parameter
  - Both methods lag behind a trend
  - One can achieve the same distribution of forecast error by setting: $\alpha = 2/(N + 1)$ or $N = (2 - \alpha)/\alpha$

Comparison of MA and ES

- Differences
  - ES carries all past history (forever!)
  - MA eliminates “bad” data after N periods
  - MA requires all N past data points to compute new forecast estimate while ES only requires last forecast and last observation of ‘demand’ to continue

Using Regression for Times Series Forecasting

- Regression Methods Can be Used When a Trend is Present.
  - Model: Dt = a + bt + εt.
- If t is scaled to 1, 2, 3, . . . , -- it becomes a number i -- then the least squares estimates for *a* and *b* can be computed as follows: (n is the number of observation we have)

$$S_{xx} = \left(\left(n^2 * (n+1) * (2n+1\right) \big/ 6\right) - \left(\left(n^2 * (n+1)^2\right) \big/ 4\right)$$

$$S_{xy} = n\sum_{i=1}^{n} iD_i - \left(\left(n * (n+1)\right) \big/ 2\right) * \sum_{i=1}^{n} D_i$$

$$a = \overline{D} - b * \left(\left(n+1\right) \big/ 2\right)$$

$$b = S_{xy} \big/ S_{xx}$$

Lets try 2-28 on page 75

Month	# Visitors	Month	# Visitors
Jan	133	Apr	640
Feb	183	May	1875
Mar	285	Jun	2550

- S_{xy}=6*(1*133+2*183+3*285+4*640+5*1875+6*2550) – (6*7/2)(133+183+285+640+1875+2550)] = 52557
- S_{xx} = [(36*7*13)/6]-[(36*49)/4)]= 105
- b = (52557/105)= 500.54
- a = 944.33 – 500.54*(6+1)/2 = -807.4

Lets try 2-28 on pg 75 (cont.)

- Forecast for July: a+b*7 = -807.4 + 500.54*7 = 2696
- Forecast for August: -807.4 + 500.54*8 = 3197
  and Continued …
- However, once we get real data for July and August, we would need to re-compute S_{xx}, S_{xy}, a and b to continue forecasting – if we wish to be accurate!

Another Method When Trend is Present

- Double exponential smoothing, of which Holt’ s method is the most common example, can also be used to forecast when there is a linear trend present in the data. The method requires separate smoothing constants for slope and intercept.
- The advantage is that once we begin building a forecast model, we can quickly revise the slope and signal constituents with the separate smoothing coefficients.

Holt’ s Method Explained

- We begin with an estimate of the intercept and slope at the start (by Lin. Reg.?)
- $S_i = \square D_i + (1-\square)(S_{i-1} + G_{i-1})$
- $G_i = \square(S_i - S_{i-1}) + (1- \square)G_{i-1}$
  - D_i is obs. demand; S_i is current est. of ‘intercept’ ;
  - G_i is current est. of slope; S_{i-1} is last est. of ‘intercept’ ; G_{i-1} is last est. of slope
- $F_{t,t+\square} = S_t + \square * G_t$ (forecast for time $\square$ into the future)

Doing earlier problem with Holts Method:

- Use **a** as 1st estimate of intercept (S_0): -807.4
- Use **b** as 1st estimate of slope (G_0): 500.54
- Working forward (we will use our data in hand to tune our model!)
- Lets set [illegible] at .15 and [illegible] at .1
- $S_{jan} = .15D_{jan} + (1-.15)(S_0 + G_0) = .15*133 + .85*(-807.4 + 500.54) = -240.87$
- $G_{jan} = .1(S_{jan} – S_0) + (1-.1)(G_0) = .1(-240.87 – (-807.4)) + .9*500.54 = 101.16+741.81 = 507.13$

Doing earlier problem with Holts Method:

- $S_{feb} = .15D_{feb} + .85*(S_{jan} + G_{jan}) = .15*183 + .85*(-240.87 + 507.13) = 253.7$
- $G_{feb} = .1(S_{feb} – S_{jan}) + .9*G_{jan} = .1(253.7 + 240.87) + .9*507.2 = 505.9$
- Continue to refine the estimate through June to build a best estimate model to forecast the future
- To update the model after demand for July is obtained we just reapply Holt’s Method and continue!

Do one with your Team:

- Try Problem 30 on page 77.

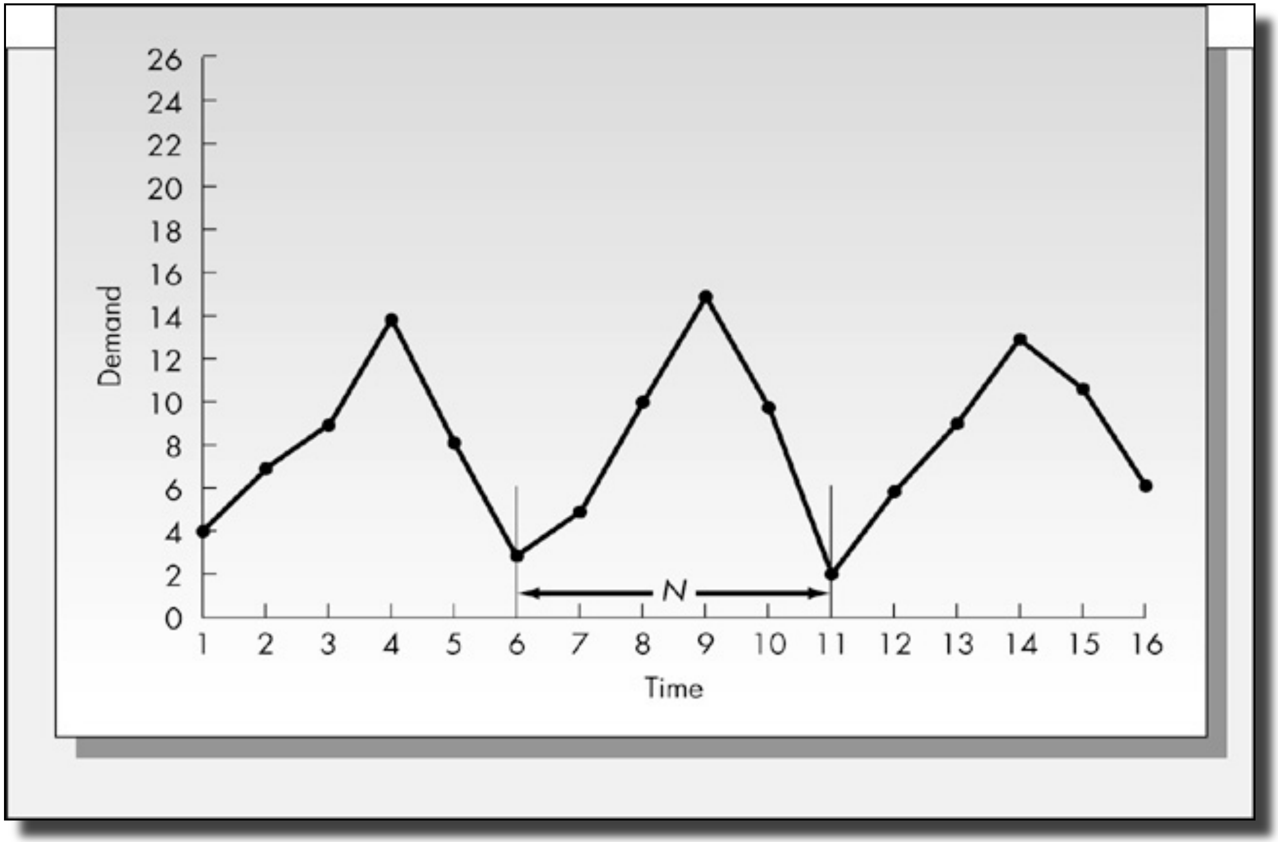
Forecasting For Seasonal Series

- Seasonality corresponds to a pattern in the data that repeats at regular intervals. (See figure next slide)
- Multiplicative seasonal factors: c_1 , c_2 , . . . , c_N where i = 1 is first season of cycle, i = 2 is second season of the cycle, etc.

$$\Sigma c_i = N$$

- c_i = 1.25 implies a 'demand' 25% higher than the baseline
- c_i = 0.75 implies 25% lower than the baseline

A Seasonal Demand Series



Quick and Dirty Method of Estimating Seasonal Factors

- Compute the sample mean of the entire data set (should be at least several cycles of data)
- Divide each observation by the sample mean (This gives a factor for each observation)
- Average the factors for like seasons
- The resulting n numbers will exactly add to N and correspond to the N seasonal factors.

Deseasonalizing a Series

- To remove seasonality from a series, simply divide each observation in the series by the appropriate seasonal factor. The resulting series will have no seasonality and may then be predicted using an appropriate method.
- Once a forecast is made on the deseasonalized series, one then multiplies that forecast by the appropriate seasonal factor to obtain a forecast for the original series.

Lets do one:

Season	Cycle	Demand	Season	Cycle	Demand
Q1	2001	205	Q1	2002	225
Q2	2001	225	Q2	2002	248
Q3	2001	185	Q3	2002	203
Q4	2001	285	Q4	2002	310

- Expected Demand $Q1,_{03}$ = (205+225)/2 = 215
- Expected Demand $Q2,_{03}$ = (225+248)/2 = 236.5
- Expected Demand $Q3,_{03}$ = (185+203)/2 = 194
- Expected Demand $Q4,_{03}$ = (285+310)/2 = 298
- Overall average demand: $\Sigma D_i/8 = 235.75$

Lets do one:

- Seasonal Demand Factors:
  - Q1: 215/235.75 = 0.912
  - Q2: 236.5/235.75 = 1.003
  - Q3: 194/235.75 = 0.823
  - Q4: 297.5/235.75 = <u>1.262</u>
  - Sum: 4.000
- The seasonal factors must sum to the number of seasons in the cycle
  - if they do not factors must be 'Normalized': $\{(\text{number of seasons in cycle})/\Sigma c_j\} * c_{ji}$ where i represents each season in the cycle

Deseasonalize the Demand

- For each period: D_i/c_j
- Work on Trends on the deseasonalized data

Period	A. Demand	Period Avg.	Period Factor	Deseason Demand (y)	Period in string (x)	X*Y	X^2
Q1 01	205	215	.912	224.78	1	224.78	1
Q2 01	225	236.5	1.003	224.29	2	448.57	4
Q3 01	185	194	.823	224.81	3	674.43	9
Q4 01	285	298	1.262	225.84	4	903.36	16
Q1 02	225	215	.912	246.72	5	1233.6	25
Q2 02	248	236.5	1.003	247.21	6	1483.26	36
Q3 02	203	194	.823	246.64	7	1726.48	49
Q4 02	310	298	1.262	245.66	8	1965.28	64
					SUM:	**8659.76**	**204**

Using Deseasonalized Tableau

- Here, X_{avg} = (1+2+3+4+5+6+7+8)/8 = 4.5
- Y_{avg} = 235.75
- b = $\boxed{\left(\sum xy - n*\bar{x}*\bar{y}\right) \Big/ \left(\sum x^2 - n*\left(\bar{x}\right)^2\right)}$
- a = $\boxed{\bar{y} - b*\bar{x}}$

Finally:

- Substituting to solve for b: 4.113
- Solving for a: 217.24
- Now we reseasonalize our forecasts for ‘03

Forecast:

Q1, ‘03	a+bx	217.24+4.113*9	254.26 *.912	231.9	232
Q2, ‘03	a+bx	217.24+4.113*10	258.37*1.003	259.2	259
Q3, ‘03	a+bx	217.24+4.113*11	262.48*.823	216.0	216
Q4, ‘03	a+bx	217.24+4.113*12	266.6*1.262	336.4	336

But what about new data?

- Same problem prevails as before updating is ‘expensive’
- As new data becomes available, we must start over to get seasonal factors, trend and intercept estimates
- Isn’t there a method to smooth this seasonalized technique?
- Yes, its called Winter’s Method or triple exponential smoothing

Exploring Winter’s Method

- This model uses 3 smoothing constants
- One for the signal, one for the trend and one for seasonal factors
- Uses this model to project the future:

$$D_t = (\mu + G * t)c_t + \varepsilon_t$$

here:

μ **is the base signal or the 'Intercept' of demand at time = zero**

G is trend (slope) component of demand

$\mathbf{c_t}$ **is seasonal component for time of interest**

ε_t **is error term**

Using Winters Methods:

- The Time Series: (deseasonalized)

$$S_t = \alpha\left(\frac{D_t}{c_{t-n}}\right) + (1-\alpha)(S_{t-1} + G_{t-1})$$

- The Trend:

$$G_t = \beta(S_t - S_{t-1}) + (1-\beta)G_{t-1}$$

- The Seasonal Factors:

$$c_t = \gamma\left(\frac{D_t}{S_t}\right) + (1-\gamma)c_{t-n}$$

where c_{t-n} was seasonal factor from the last cycle of data

Beginning Winter’s Method:

- We must derive initial estimates of the 3 values: S_t, G_t and c_t’s
- Typically we set: [illegible] = 2[illegible] = 2[illegible]
- Deriving initial estimates takes at least two complete cycles of data

Compute Slope Estimate:

- Compute sample means for each cycle of data (V_1 and V_2)

$$V_1 = \frac{1}{n} \sum_{j=-2n+1}^{-n} D_j$$

average demand for 2 cycles 'ago'

$$V_2 = \frac{1}{n} \sum_{j=-n+1}^{0} D_j$$

Avg. Demand in last cycle

here: n is # of seasons in cycle

j = 0 is today!

- Slope estimate is:

$$G_0 = \frac{(V_2 - V_1)}{n}$$

Signal and seasonal factor estimates:

- Signal:

$$S_0 = V_2 + G_0 \left[{(n-1)} \big/ {2} \right]$$

- Seasonal Factor Estimates:

$$c_t = {D_t} \Big/ {\left[V_i - \left(\left(\frac{n+1}{2} \right) - j \right) * G_0 \right]}$$

i = 1 for 1st cycle; 2 for last cycle

j is season in cycle

(j = 1 for t = -2n+1 & -n+1, etc.)

Compute Seasonal Factor Averages & Normalize

- Averaging c_i:

$$c_{-n+1} = \left(\left(c_{-2n+1} + c_{-n+1} \right) \Big/ 2 \right)$$

1st season of cycle

K

$$c_0 = \left(\left(c_{-n} + c_0 \right) \Big/ 2 \right)$$

last season of cycle

- Normalizing:

$$c_j = \left[c_j \Big/ \sum_{i=-n+1}^{0} c_i \right] * n$$

denominator is sum of factors

We have the Model so lets do one:

Period	Obs. Demand	Period	Obs. Demand
1st Q ('01)	72	1st Q ('02)	83
2nd Q ('01)	107	2nd Q ('02)	121
3rd Q ('01)	55	3rd Q ('02)	63
4th Q ('01)	88	4th Q ('02)	100

Initial Slope & Intercept Estimates:

$$V_1 = \left(\tfrac{1}{4} \right) * \left(72 + 107 + 55 + 88 \right) = 80.5$$

$$V_2 = \left(\tfrac{1}{4} \right) * \left(83 + 121 + 63 + 100 \right) = 91.75$$

$$\therefore$$

$$G_0 = \left(\frac{V_2 - V_1}{4} \right) = \left(\frac{(91.75 - 80.5}{4} \right) = 2.813$$

$$S_0 = V_2 + G_0 \left(\frac{(4 - 1)}{2} \right) = 95.97$$

The 4 Seasonal Factors (1st season – in each cycle)

$$c_t = \frac{D_t}{\left[V_i - \left(\left(\frac{(n+1)}{2}\right) - j\right) * G_0\right]}$$

1st season (n=-7; n=-3); here j = 1

$$c_{-7} = \frac{72}{\left(80.5 - \left[\frac{5}{2} - 1\right] * 2.813\right)} = .944$$

$$c_{-3} = \frac{83}{\left(91.75 - \left[\frac{5}{2} - 1\right] * 2.813\right)} = .948$$

$$avg : \frac{(.944 + .948)}{2} = .946$$

Same approach: Seasons 2, 3 & 4

- 2nd Season: c_{-6} & c_{-2} $(j = 2)$ 1.353; 1.340: avg is 1.346
- 3rd Season: c_{-5} & c_{-1} $(j = 3)$ .671; .676: avg is 0.674
- 4th Season: c_{-4} & c_{0} $(j = 4)$ 1.039; 1.042: avg is 1.040
- c_i = .946 + 1.346 + .676 + 1.040 = 4.008
  so may want to (should?) normalize!

Normalizing:

$$c_j = \left[\frac{n}{\sum_{i=0}^{-n+1} c_i} \right] * c_j$$

denominator is sum of factors

- Here:
  - $c_{1st} = (4/4.008)*.946 = .944$
  - $c_{2nd} = (4/4.008)*1.346 = 1.343$
  - $c_{3rd} = (4/4.008)*.676 = .675$
  - $c_{4th} = (4/4.008)*1.040 = 1.038$

Making a Forecast:

$$F_{t,t+\tau} = \left(S_t + \tau G_t\right) c_{(t+\tau)-n}$$

$$F_{0,1stQ} = \left(S_0 + 1 * G_0\right) c_{1stQ} =$$

$$\left(95.97 + 1 * 2.813\right) * .944 = 93.25 \rightarrow (94)$$

$$F_{0,2ndQ} = \left(95.97 + 2 * 2.813\right) * 1.343 = 136.44$$

$$F_{0,3rdQ} = \left(95.97 + 3 * 2.813\right) * .675 = 70.48$$

$$F_{0,4thQ} = \left(95.97 + 4 * 2.813\right) * 1.038 = 111.30$$

Updating Model (if D_1 is 112)

$$S_1 = \alpha\left(\frac{D_1}{c_{1stQ}}\right) + (1-\alpha)(S_0 + G_0) =$$

$$.2\left(\frac{112}{.944}\right) + (.8) * (95.97 + 2.813) = 102.76$$

$$G_1 = \beta(S_1 - S_0) + (1-\beta)G_0$$

$$= .1 * (102.76 - 95.97) + (.9) * 2.813 = 3.210$$

$$c_{1stQ} = \gamma\left(\frac{D_1}{S_1}\right) + (1-\gamma)c_{1stQ}$$

$$= .1 * \left(\frac{112}{102.76}\right) + .9 * .944 = .959$$

With updated c_{1stQ} we **may** have to renormalize all c_i's

Fine-tuning the Model

- To adequately employ Winter’s method, we need 3 cycles of historic data
- Two cycles for model development
- 3rd cycle for fine tuning by update method
- Forecast can then begin after adjustments
- Update each term (S,G and c_i’s as new facts roll in)

Lets Try 1:

- Prob 35 & 36 on page 88 in text.

Practical Considerations

- Overly sophisticated forecasting methods can be problematic, especially for long term forecasting. (Refer to Figure on the next slide.)

- Tracking signals may be useful for indicating forecast bias.

- Box-Jenkins methods require substantial data history, use the correlation structure of the data, and can provide significantly improved forecasts under some circumstances.

The Difficulty with Long-Term Forecasts –an afterthought!