ÖZYEĞİN ÜNİVERSİTESİ

M7
Predictive Analytics

ENİS KAYIŞ

Moving Averages and Smoothing Methods

- We will describe three simple approaches to forecasting a time series:
  - Naïve
  - Averaging
  - Smoothing
- Naive methods are used to develop simple models that assume that very recent data provide the best predictors of the future.
- Averaging methods generate forecasts based on an average of past observations.
- Smoothing methods produce forecasts by averaging past values of a series with a decreasing (exponential) series of weights.

Moving Averages and Smoothing Methods

Forecasting Procedure

- You observe Y_t at time t
- Then use the past data $Y_{t-1}, Y_{t-2}, Y_{t-3}$
- Obtain forecasts for different periods in the future

Past data	You are here t	Periods to be forecast
$Y_{t-3}, Y_{t-2}, Y_{t-1},$	$Y_t,$	$\hat{Y}_{t+1}, \hat{Y}_{t+2}, \hat{Y}_{t+3}, \ldots$

Y_t is the most recent observation of a variable

$\hat{Y}_{t+1}$ is the forecast for one period in the future

Moving Averages and Smoothing Methods

- Steps for Evaluating Forecasting Methods
  - Select an appropriate forecasting method based on the nature of the data and experience.
  - Divide the observations into two parts.
  - Apply the selected forecasting method to develop fitted model for the first part of the data.
  - Use the fitted model to forecast the second part of the data.
  - Calculate the forecasting errors and evaluate using measures of accuracy.
  - Make your decision to accept or modify or reject the suggested method or develop another method to compare the results.

Moving Averages and Smoothing Methods

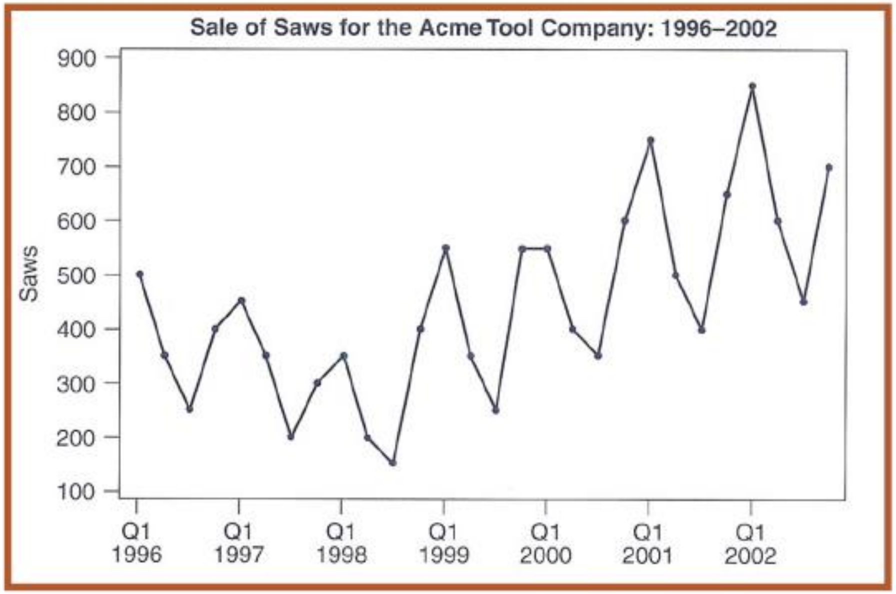
- **Naïve Methods**
- These methods assume that the most recent information is the best predictors of the future.
- For the stationary data, we use the simplest model is $\hat{Y}_{t+1} = Y_t$.
- For the nonstationary (trended) data, we use the following models
  - The amount of change model is $\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1})$.
  - The rate of change model is $\hat{Y}_{t+1} = Y_t \frac{\mathrm{Y_t}}{Y_{t-1}}$.
- For the strongly seasonal (quarterly) data, we use the following models
  - The quarterly data model is $\hat{Y}_{t+1} = Y_{t-3}$.
  - The monthly data model is $\hat{Y}_{t+1} = Y_{t-11}$

Moving Averages and Smoothing Methods

- For the seasonal (quarterly) and trended data, we use

$$\hat{Y}_{t+1} = Y_{t-3} + 4\frac{Y_t - Y_{t-4}}{4} = Y_{t-3} + Y_t - Y_{t-4}$$

- Example 1:
  - The figure shows the quarterly sales of saws for the Acme Tool Company. The naive technique forecasts that sales for the next quarter will be the same as the previous quarter.



Moving Averages and Smoothing Methods

- The data from 1996 to 2002

Year	*Quarter*	*t*	*Sales*	*Year*	*Quarter*	*t*	*Sales*
1996	1	1	500		3	15	250
	2	2	350		4	16	550
	3	3	250	2000	1	17	550
	4	4	400		2	18	400
1997	1	5	450		3	19	350
	2	6	350		4	20	600
	3	7	200	2001	1	21	750
	4	8	300		2	22	500
1998	1	9	350		3	23	400
	2	10	200		4	24	650
	3	11	150	2002	1	25	850
	4	12	400		2	26	600
1999	1	13	550		3	27	450
	2	14	350		4	28	700

- If the data from 1996 to 2001 are used as the initialization part and 2002 as the test part.
- The forecast for the first quarter of 2002

$$\hat{Y}_{24+1} = Y_{24}$$
$$\hat{Y}_{25} = 650$$

Moving Averages and Smoothing Methods

- The forecasting error is $e_{25} = Y_{25} - \hat{Y}_{25} = 850 - 650 = 200$
- Similarly, the forecast for period 26 is 850 with an error of -250.
- But its very clear that the data have an upward trend and there appears to be a seasonal pattern (first and fourth quarters are relatively high).
- A decision must be made to modify the naive model!
- Our technique can be adjusted to take trend into consideration by adding the difference between this period and the last period.
- The forecast equation is $\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1})$
- The forecast for the first quarter of 2002 is

$$\begin{aligned}
\hat{Y}_{24+1} &= Y_{24} + (Y_{24} - Y_{24-1}) \\
\hat{Y}_{25} &= Y_{24} + (Y_{24} - Y_{23}) \\
\hat{Y}_{25} &= 650 + (650 - 400) \\
\hat{Y}_{25} &= 650 + 250 = 900
\end{aligned}$$

Moving Averages and Smoothing Methods

- The forecasting error is $e_{25} = Y_{25} - \hat{Y}_{25} = 850 - 900 = -50$
- Visual inspection of the data in indicates that seasonal variation seems to exist.
- Sales in the first and fourth quarters are typically larger than those in any of the other quarters.
- If the seasonal pattern is strong, an appropriate forecast equation for quarterly data might be $\hat{Y}_{t+1} = Y_{t-3}$ $\boxed{1}$
- The previous equation says that next quarter the variable will take on the same value that it did in the corresponding quarter one year ago.
- The major weakness of this approach is that it ignores everything that has occurred since last year and also any trend

Moving Averages and Smoothing Methods

- There are several ways of introducing more recent information.
- For example, the analyst can combine seasonal and trend estimates and forecast the next quarter using

$$\boxed{\hat{Y}_{t+1} = Y_{t-3} + 4\frac{Y_t - Y_{t-4}}{4} = Y_{t-3} + Y_t - Y_{t-4}} \quad \boxed{2}$$

- The Y_{t-3} term forecasts the seasonal pattern, and the remaining term averages the amount of change for the past four quarters and provides an estimate of the trend.
- The naive forecasting models in the last equations are given for quarterly data.
- Adjustments can be made for data collected over different time periods.
- For monthly data, the seasonal period is 12, not 4, and the forecast for the next period (month) given by

$$\boxed{\hat{Y}_{t+1} = Y_{t-11}}$$

Moving Averages and Smoothing Methods

- For some purposes, the rate of change might be more appropriate than the absolute amount of change. If so, it is reasonable to generate forecasts according to

$$\hat{Y}_{t+1} = Y_t \frac{Y_t}{Y_{t-1}} \qquad \boxed{3}$$

- The forecasts for the first quarter of 2002 using Equations 1, 2 and 3 are

$$\hat{Y}_{24+1} = Y_{24} \frac{Y_{24}}{Y_{24-1}} = Y_{24} \frac{Y_{24}}{Y_{23}}$$

$$\hat{Y}_{25} = 650 \frac{650}{400} = 1{,}056$$

$$\hat{Y}_{24+1} = Y_{24-3} = Y_{21}$$

$$\hat{Y}_{25} = Y_{21} = 750$$

$$\hat{Y}_{24+1} = Y_{24-3} + 4\frac{Y_{24} - Y_{24-4}}{4} = Y_{24-3} + Y_{24} - Y_{24-4}$$

$$\hat{Y}_{25} = Y_{21} + Y_{24} - Y_{20}$$

$$\hat{Y}_{25} = 750 + 650 - 600 = 800$$

Moving Averages and Smoothing Methods

Averaging methods

- Frequently, management faces the situation in which forecasts need to be updated daily, weekly, or monthly for inventories containing hundreds or thousands of items.
- Often it is not possible to develop sophisticated forecasting techniques for each item.
- Instead, some quick, inexpensive, very simple short-term forecasting tools are needed to accomplish this task.
- These tools assume that the most recent information is used to predict the future and the forecasts need to be updated daily, weekly, monthly, etc.
- They use a form of weighted average of past observations to smooth short-term fluctuations.
- In this case different averaging methods should be considered.

Moving Averages and Smoothing Methods

Simple Average method

- This method uses the mean of all relevant historical observations as the forecast of the next period.
- Also, it used when the forces generating series have stabilized and the environment in which the series exists is generally unchanging.
- The objective is to use past data to develop a forecasting model for future periods.
- As with the naive methods, a decision is made to use the first *t* data points as the initialization part and the remaining data as a test part.
- Next, the following equation is used to average the initialization part of the data and to forecast the next period.

$$\boxed{\hat{Y}_{t+1} = \frac{1}{t}\sum_{i=1}^{t} Y_i}$$

- When a new observation becomes available, the forecast for the next period ($t + 2$) is the average of the above equation and that data point

Moving Averages and Smoothing Methods

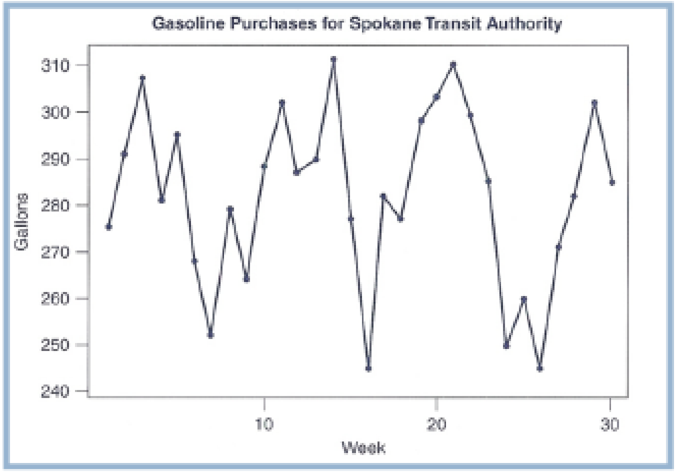
Example: 2

The Spokane Transit Authority (STA) operates a fleet of vans used to transport both the disabled and elderly. A record of the gasoline purchased for this fleet of vans is shown in table below. The actual amount of gasoline consumed by a van on a given day is determined by the random nature of the calls and the destinations.

Week t	Gallons Y_t	Week t	Gallons Y_t	Week t	Gallons Y_t
1	275	11	302	21	310
2	291	12	287	22	299
3	307	13	290	23	285
4	281	14	311	24	250
5	295	15	277	25	260
6	268	16	245	26	245
7	252	17	282	27	271
8	279	18	277	28	282
9	264	19	298	29	302
10	288	20	303	30	285

Moving Averages and Smoothing Methods

- Examination of the gasoline purchases plotted in following figure shows the data are very stable.



- Hence, the method of simple averages is used for weeks 1 to 28 to forecast gasoline purchases for weeks 29 and 30.

- The forecast for week 29 is

$$\hat{Y}_{28+1} = \frac{1}{28}\sum_{i=1}^{28} Y_i$$

$$\hat{Y}_{29} = \frac{7{,}874}{28} = 281.2$$

Moving Averages and Smoothing Methods

- The forecasting error is

$$e_{29} = Y_{29} - \hat{Y}_{29} = 302 - 281.2 = 20.8$$

- The forecast for week 30 includes one more data point (302) added to the initialization period.
- So the forecast is

$$\hat{Y}_{28+2} = \frac{28\hat{Y}_{28+1} + Y_{28+1}}{28+1} = \frac{28\hat{Y}_{29} + Y_{29}}{29}$$

$$\hat{Y}_{30} = \frac{28(281.2) + 302}{29} = 281.9$$

- The forecasting error is

$$e_{30} = Y_{30} - \hat{Y}_{30} = 285 - 281.9 = 3.1$$

- Using the method of simple averages, the forecast of gallons of gasoline purchased for week 31 is

$$\hat{Y}_{30+1} = \frac{1}{30}\sum_{i=1}^{30} Y_i = \frac{8,461}{30} = 282$$

Moving Averages and Smoothing Methods

Moving Averages

- The method of simple averages uses the mean of all the data to forecast. What if the analyst is more concerned with recent observations?
- A constant number of data points can be specified at the outset and a mean computed for the most recent observations.
- A *moving average* of order k is the mean value of k consecutive observations. The most recent moving average value provides a forecast for the next period.
- The term moving average is used to describe this approach.
- As each new observation becomes available, a new mean is computed by adding the newest value and dropping the oldest.
- This moving average is then used to forecast the next period.
- The simple moving average forecast of the order k, MA(k), is given by

$$\boxed{\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-2} + \cdots + Y_{t-k+1}}{k}}$$

Moving Averages and Smoothing Methods

Moving Averages

- Moving average models do not handle the trend and the seasonality very well, It is useful in stationary time series data.
- The method of simple averages uses the mean of all the data to forecast. What if the analyst is more concerned with recent observations?
- In the moving average, equal weights are assigned to each observation.
- Each new data point is included in the average as it becomes available, and the earliest data point is discarded.
- The rate of response to changes in the underlying data pattern depends on the number of periods, *k*, included in the moving average.
- A moving average of order 1, MA(1), is simply the naive forecasting.

Moving Averages and Smoothing Methods

ÖZYEĞİN UNIVERSITY

Example: 3

The table below demonstrates the moving average forecasting technique with the Spokane Transit Authority data, using a five-week moving average.

t	Gallons	$\hat{Y}_t$	e_t	t	Gallons	$\hat{Y}_t$	e_t
1	275	—	—	16	245	293.4	−48.4
2	291	—	—	17	282	282.0	0.0
3	307	—	—	18	277	281.0	−4.0
4	281	—	—	19	298	278.4	19.6
5	295	—	—	20	303	275.8	27.2
6	268	289.8	−21.8	21	310	281.0	29.0
7	252	288.4	−36.4	22	299	294.0	5.0
8	279	280.6	−1.6	23	285	297.4	−12.4
9	264	275.0	−11.0	24	250	299.0	−49.0
10	288	271.6	16.4	25	260	289.4	−29.4
11	302	270.2	31.8	26	245	280.8	−35.8
12	287	277.0	10.0	27	271	267.8	3.2
13	290	284.0	6.0	28	282	262.2	19.8
14	311	286.2	24.8	29	302	261.6	40.4
15	277	295.6	−18.6	30	285	272.0	13.0

Moving Averages and Smoothing Methods

- The moving average forecast for week 29 is

$$\hat{Y}_{28+1} = \frac{Y_{28} + Y_{28-1} + \cdots + Y_{28-5+1}}{5}$$

$$\hat{Y}_{29} = \frac{Y_{28} + Y_{27} + Y_{26} + Y_{25} + Y_{24}}{5}$$

$$\hat{Y}_{29} = \frac{282 + 271 + 245 + 260 + 250}{5} = \frac{1{,}308}{5} = 261.6$$

- The forecasting error is

$$e_{29} = Y_{29} - \hat{Y}_{29} = 302 - 261.6 = 40.4$$

- The forecast for week 31 is

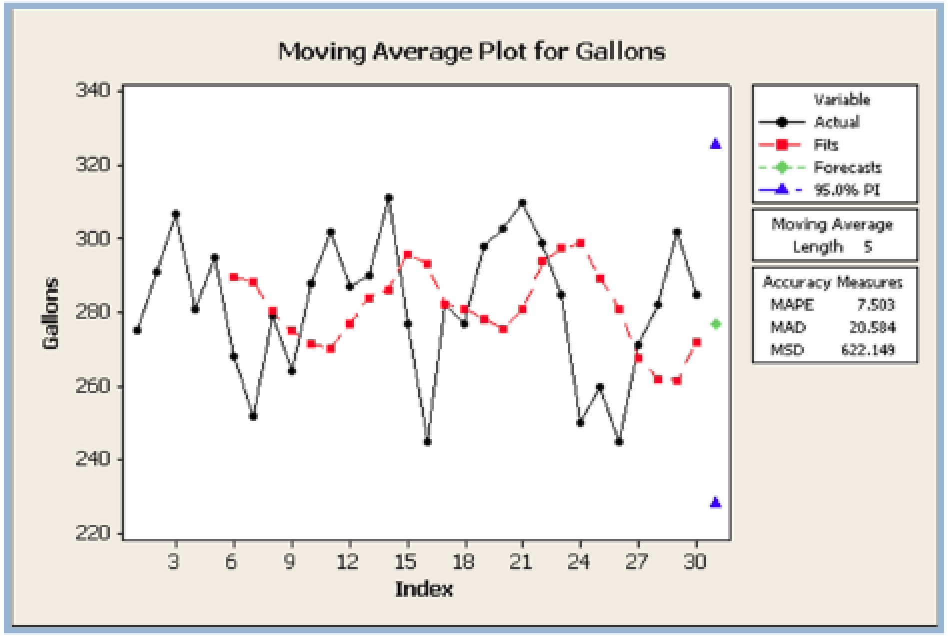
$$\hat{Y}_{30+1} = \frac{Y_{30} + Y_{30-1} + \cdots + Y_{30-5+1}}{5}$$

$$\hat{Y}_{31} = \frac{Y_{30} + Y_{29} + Y_{28} + Y_{27} + Y_{26}}{5}$$

$$\hat{Y}_{31} = \frac{285 + 302 + 282 + 271 + 245}{5} = \frac{1{,}385}{5} = 277$$

Moving Averages and Smoothing Methods

- The following figure shows the five-week moving average plotted against the actual data.



Moving Averages and Smoothing Methods

- The analyst must use judgment when determining how many days, weeks, months, or quarters on which to base the moving average.
- The smaller the number, the more weight is given to recent periods. Conversely, the greater the number, the less weight is given to more recent periods.
- A small number is most desirable when there are sudden shifts in the level of the series.
- A small number places heavy weight on recent history, which enables the forecasts to catch up more rapidly to the current level.
- A large number is desirable when there are wide, infrequent fluctuations in the series.

Moving Averages and Smoothing Methods

- For quarterly data, a four-quarter moving average, MA(4), yields an average of the four quarters, and for monthly data, a 12-month moving average, MA(12), eliminates or averages out seasonal effects.
- The larger the order of the moving average, the greater the smoothing effect.

Moving Averages and Smoothing Methods

Double Moving Averages

This method used when the time series data are nonstationary or trended and this happen by computing the moving average of order k to the first set of moving averages.

Let

$$M_t = \hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-2} + \cdots + Y_{t-k+1}}{k}$$

First we compute the moving average of order *k*.

Then

$$M'_t = \frac{M_t + M_{t-1} + M_{t-2} + \cdots + M_{t-k+1}}{k}$$

Then we compute the second moving average of order *k*.

Thus the Model forecast m period is given by

$$\hat{Y}_{t+p} = a_t + b_t p$$

where

$$a_t = M_t + (M_t - M'_t) = 2M_t - M'_t$$

$$b_t = \frac{2}{k-1}(M_t - M'_t)$$

Remember differencing for trended time series

k = number of periods in the moving average

p = number of periods ahead to be forecast

Moving Averages and Smoothing Methods

Example 4:

- In a particular city, Camping Gear offers rental tents for customers. The company is growing and needs to expand its inventory to accommodate the increasing demand for its services.
- The management wants to forecast rentals for the next month.
- At first, Jill attempts to develop a three-week moving average.
- The MSE for this model is 133.
- Because the data are obviously trending, she finds that her forecasts are consistently underestimating actual rentals.

t	Weekly Units Rented Y_t	Three-Week Moving Total	Moving Average Forecast $\hat{Y}_{t+1}$	e_t
1	654	—	—	—
2	658	—	—	—
3	665	1,977	—	—
4	672	1,995	659	13
5	673	2,010	665	8
6	671	2,016	670	1
7	693	2,037	672	21
8	694	2,058	679	15
9	701	2,088	686	15
10	703	2,098	696	7
11	702	2,106	699	3
12	710	2,115	702	8
13	712	2,124	705	7
14	711	2,133	708	3
15	728	2,151	711	17
16	—	—	717	—

$MSE = 133.$

Moving Averages and Smoothing Methods

- The results for double moving average are shown in the table below

(1) Time t	(2) Weekly Sales Y_t	(3) Three-Week Moving Average M_t	(4) Double Moving Average, M'_t	(5) Value of a	(6) Value of b	(7) Forecast $a + bp$ $(p = 1)$	(8) e_t
1	654	—	—	—	—	—	—
2	658	—	—	—	—	—	—
3	665	659	—	—	—	—	—
4	672	665	—	—	—	—	—
5	673	670	665	675	5	—	—
6	671	672	669	675	3	680	−9
7	693	679	674	684	5	678	15
8	694	686	679	693	7	689	5
9	701	696	687	705	9	700	1
10	703	699	694	704	5	714	−11
11	702	702	699	705	3	709	−7
12	710	705	702	708	3	708	2
13	712	708	705	711	3	711	1
14	711	711	708	714	3	714	−3
15	728	717	712	722	5	717	11
16	—	—	—	—	—	727	—

$MSE = 63.7.$

Moving Averages and Smoothing Methods

- To understand the forecast for *week 16*, the computations are:
- First we computed the three week moving average (column 3) using

$$M_{15} = \hat{Y}_{15+1} = \frac{Y_{15} + Y_{15-1} + Y_{15-3+1}}{3}$$

$$M_{15} = \hat{Y}_{16} = \frac{728 + 711 + 712}{3} = 717$$

- To compute the double moving averages (column 4), we used

$$M'_{15} = \frac{M_{15} + M_{15-1} + M_{15-3+1}}{3}$$

$$M'_{15} = \frac{717 + 711 + 708}{3} = 712$$

- The difference between the two moving averages (column 5) is computed

$$a_{15} = 2M_{15} - M'_{15} = 2(717) - 712 = 722$$

- The slope is adjusted by

$$b_{15} = \frac{2}{3-1}(M_{15} - M'_{15}) = \frac{2}{2}(717 - 712) = 5$$

Moving Averages and Smoothing Methods

- The forecast is done for one period into the future (column 7).

$$\hat{Y}_{15+1} = a_{15} + b_{15}p = 722 + 5(1) = 727$$

- The forecast four weeks into the future is

$$\hat{Y}_{15+4} = a_{15} + b_{15}p = 722 + 5(4) = 742$$

Note that the MSE has been reduced from 133 to 63.7.

- It seems reasonable that more recent observations are likely to contain more important information.
- A procedure is introduced in the next section that gives more emphasis to the most recent observations.

Moving Averages and Smoothing Methods

Smoothing Methods

- These methods also assume that the most recent information is used to predict the future and the forecasts need to be updated daily, weekly, monthly, etc.
- In this case different smoothing methods should be considered.

1. *Exponential Smoothing*

- This method continually revises a forecast in the light of more recent experiences.
- It is often appropriate for data with no predictable upward or downward trend.
- It provides an exponentially weighted moving average of all previously observed values.
- The most recent observation receives the largest weight α (where $0 < \alpha < 1$).

Moving Averages and Smoothing Methods

- The next most recent observation receives less weight, $\alpha(1 - \alpha)$, the observation two time periods in the past even less weight, $\alpha(1 - \alpha)^2$, and so forth.

- The weight c (where $0 < \alpha < 1$) is given to the newly observed value, and weight $(1 - \alpha)$ is given to the old forecast.

New forecast = [α x (new observation)] + [(1 — α) x (old forecast)]

- More formally, the exponential smoothing equation is

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

where

$\hat{Y}_{t+1}$ = new smoothed value or the forecast value for the next period
α = smoothing constant ($0 < \alpha < 1$)
Y_t = new observation or actual value of series in period t
$\hat{Y}_t$ = old smoothed value or forecast for period t

- Also, it can be rewritten as

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

Moving Averages and Smoothing Methods

- Where α is called smoothing constant or the weighting is factor and $\hat{Y}_t$ is the old smoothed value or forecast with $\hat{Y}_1 = Y_1$.
- Note that $\hat{Y}_{t+1}$ is recognized as an exponentially smoothed value since mathematically it is equivalent to

$$\boxed{\hat{Y}_{t+1} = \sum_{m=0}^{\infty} \alpha(1-\alpha)^m Y_{t-m}}$$

or

$$\boxed{\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + \cdots}$$

- The value assigned to α is the key to the analysis.
- If it is desired that predictions be stable and random variations smoothed, a small value of α is required.
- If a rapid response to a real change in the pattern of observations is desired, a larger value of α is appropriate.
- One method of estimating α is an iterative procedure that minimizes the mean squared error (*MSE*).
- Forecasts are computed for, say, α equal to .1, .2,..., .9, and the sum of the squared forecast errors is computed for each.

Moving Averages and Smoothing Methods

- The value of α producing the smallest error is chosen for use in generating future forecasts.
- To start the algorithm for, an initial value for the old smoothed series must be set.
  - One approach is to set the first smoothed value equal to the first observation.
  - Another method is to use the average of the first five or six observations for the initial smoothed value.
- Exponential smoothing is often a good forecasting method when a nonrandom time series exhibits trending behavior.

Moving Averages and Smoothing Methods

Example: 5

The exponential smoothing technique is demonstrated in the following table and figure for Acme Tool Company for the years 1996 to 2002, using smoothing constants of .1 and .6. The data for the first quarter of 2002 will be used as test data to help determine the best value of α (among the two considered).

	$\alpha = .1$		$\alpha = .6$	
Period	*Calculations*	*Weight*	*Calculations*	*Weight*
t		.100		.600
$t-1$	$.9 \times .1$	.090	$.4 \times .6$	.240
$t-2$	$.9 \times .9 \times .1$	.081	$.4 \times .4 \times .6$	.096
$t-3$	$.9 \times .9 \times .9 \times .1$	.073	$.4 \times .4 \times .4 \times .6$	.038
$t-4$	$.9 \times .9 \times .9 \times .9 \times .1$	.066	$.4 \times .4 \times .4 \times .4 \times .6$	.015
All others		.590		.011
	Totals	1.000		1.000

Moving Averages and Smoothing Methods

The exponentially smoothed series is computed by initially setting $\hat{Y}_1$ equal to 500.
If earlier data are available, it might be possible to use them to develop a smoothed series up to 1996 and use this experience as the initial value for the smoothed series.

The computations leading to the forecast for periods 3 and 4 are demonstrated next.

Time: Year	Time: Quarters	Actual Value Y_t	Smoothed Value $\hat{Y}_t (\alpha = .1)$	Forecast Error e_t	Smoothed Value $\hat{Y}_t (\alpha = .6)$	Forecast Error e_t
1996	1	500	500.0	0.0	500.0	0.0
	2	350	500.0	−150.0	500.0	−150.0
	3	250	485.0 (1)	−235.0 (2)	410.0	−160.0
	4	400	461.5 (3)	−61.5	314.0	86.0
1997	5	450	455.4	−5.4	365.6	84.4
	6	350	454.8	−104.8	416.2	−66.2
	7	200	444.3	−244.3	376.5	−176.5
	8	300	419.9	−119.9	270.6	29.4
1998	9	350	407.9	−57.9	288.2	61.8
	10	200	402.1	−202.1	325.3	−125.3
	11	150	381.9	−231.9	250.1	−100.1
	12	400	358.7	41.3	190.0	210.0
1999	13	550	362.8	187.2	316.0	234.0
	14	350	381.6	−31.5	456.4	−106.4
	15	250	378.4	−128.4	392.6	−142.6
	16	550	365.6	184.4	307.0	243.0
2000	17	550	384.0	166.0	452.8	97.2
	18	400	400.6	−0.6	511.1	−111.1
	19	350	400.5	−50.5	444.5	−94.5
	20	600	395.5	204.5	387.8	212.2
2001	21	750	415.9	334.1	515.1	234.9
	22	500	449.3	−50.7	656.0	−156.0
	23	400	454.4	−54.4	562.4	−162.4
	24	650	449.0	201.0	465.0	185.0
2002	25	850	469.0	381.0	576.0	274.0

Moving Averages and Smoothing Methods

- At time period 2, the forecast for period 3 with α = .1 is

$$\hat{Y}_{2+1} = \alpha Y_2 + (1-\alpha)\hat{Y}_2$$
$$\hat{Y}_3 = .1(350) + .9(500) = 485$$

- The error in this forecast is

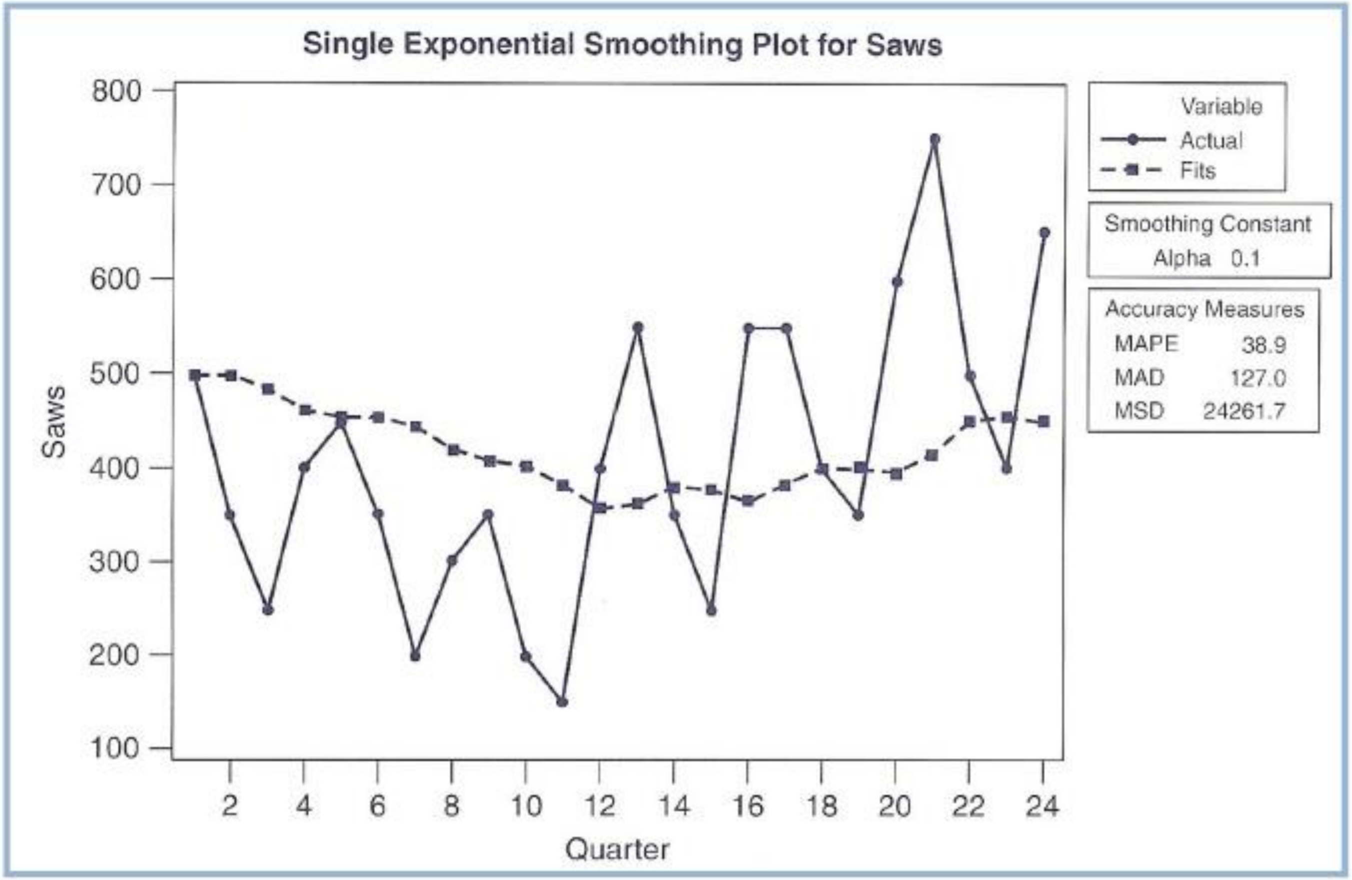
$$e_3 = Y_3 - \hat{Y}_3 = 250 - 485 = -235$$

- The forecast for period 4 is

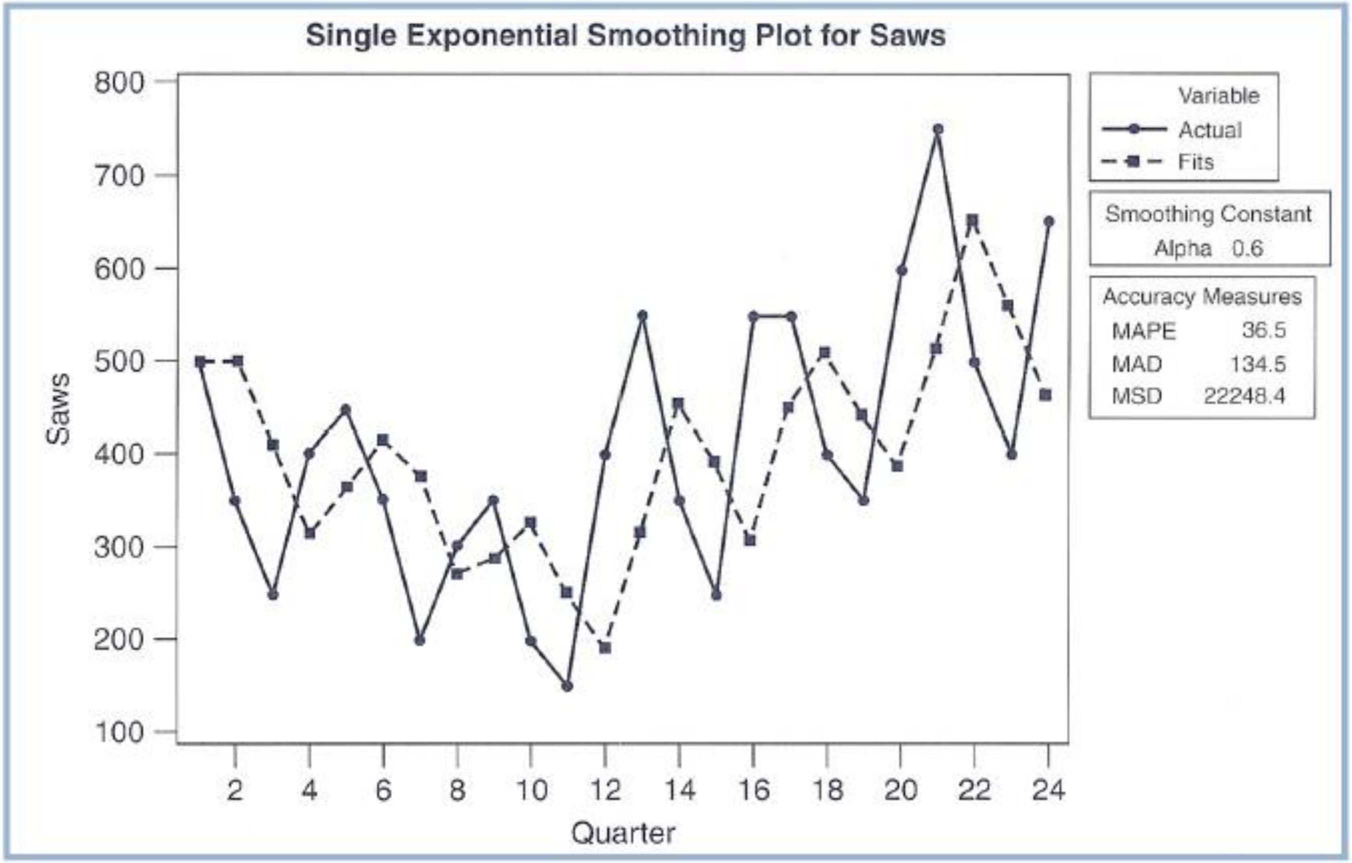
$$\hat{Y}_{3+1} = \alpha Y_3 + (1-\alpha)\hat{Y}_3$$
$$\hat{Y}_4 = .1(250) + .9(485) = 461.5$$

- From the previous table, using smoothing constant of .1, the forecast for the first quarter of 2002 is 469 with a squared error of 145,161.
- When the smoothing constant is .6, the forecast for the first quarter of 2002 is 576 with a squared error of 75,076.
- On the basis of this limited evidence, exponential smoothing with α = .6 performs better than exponential smoothing with α = .1.

Moving Averages and Smoothing Methods



Moving Averages and Smoothing Methods



Moving Averages and Smoothing Methods

- In summary

$\alpha = .1$	$MSE = 24{,}262$	$MAPE = 38.9\%$
$\alpha = .6$	$MSE = 22{,}248$	$MAPE = 36.5\%$

- *MSE* and *MAPE* are both large and, on the basis of these summary statistics.
- It is apparent that exponential smoothing does not represent these data well.
- As we shall see, a smoothing method that allows for seasonality does a better job of predicting the data in the previous example.
- Another factor beside α that affects the values of subsequent forecasts is the choice of the initial value $\hat{Y}_1$ for the smoothed series.
- In Example 4.5, $\hat{Y}_1 = Y_1$ was used as the initial smoothed value.
- This choice gave Y_1 too much weight in later forecasts.
- Fortunately, the influence of the initial forecast diminishes greatly as t increases.

Moving Averages and Smoothing Methods

- Another approach to initialize the smoothing procedure is to average the first k observations. The smoothing then begins with

$$\hat{Y}_1 = \frac{1}{k}\sum_{t=1}^{k} Y_t$$

- Often k is chosen to be a relatively small number.

Example: 6

The computation of the initial value as an average for the Acme Tool Company data presented in Example 4.5 is shown next. If k is chosen to equal 6, then the initial value is

$$\hat{Y}_1 = \frac{1}{6}\sum_{t=1}^{6} Y_t = \frac{1}{6}(500 + 350 + 250 + 400 + 450 + 350) = 383.3$$

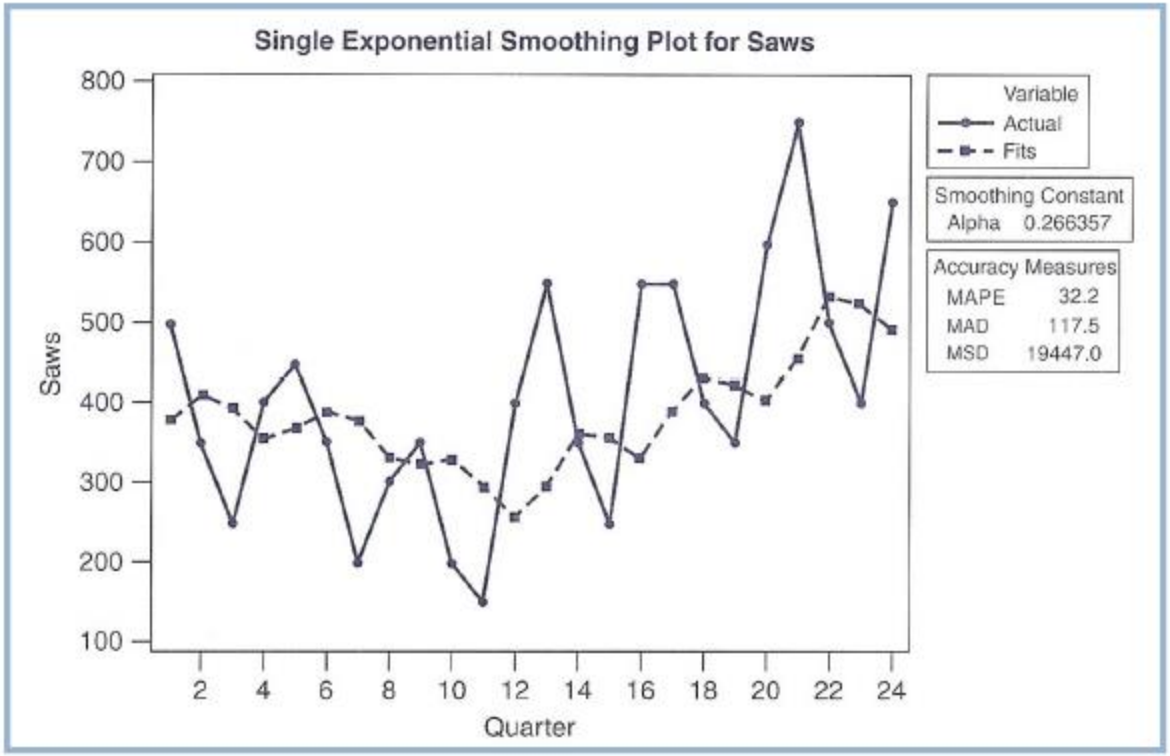
The *MSE* and *MAPE* for each alpha when an initial smoothed value of 383.3 is used are shown next.

$\alpha = .1$	$MSE = 21{,}091$	$MAPE = 32.1\%$
$\alpha = .6$	$MSE = 22{,}152$	$MAPE = 36.7\%$

Moving Averages and Smoothing Methods

- The initial value of 383.3 led to a decrease in *MSE* and *MAPE* for α = .1 but did not have much effect when α = .6.
- Now the best model, based on the *MSE* and *MAPE* summary measures, appears to be one that uses α = .1 instead of .6.
- the smoothing constant of α = .266 was automatically selected by minimizing *MSE*.
- The *MSE* is reduced to 19,447, the MAPE equals 32.2% and, although not shown, the MPE equals to - 6.4% and the forecast for the first quarter of 2002 is 534.
- The autocorrelation function for the residuals of the exponential smoothing method using an alpha of .266.
- When the Ljung-Box test is conducted for six time lags, the large value of LBQ (33.86) shows that the first six residual autocorrelations as a group are larger than would be expected if the residuals were random.

Moving Averages and Smoothing Methods



- Optimization
  - Need to choose value for the smoothing constant α and initial value l_0.
- Similarly to regression—we choose α and l_0 by minimizing SSE:

$$\text{SSE} = \sum_{t=1}^{T}(y_t - \hat{y}_{t|t-1})^2.$$

- Unlike regression there is no closed form solution—use numerical optimization.

Moving Averages and Smoothing Methods

<u>Example: 8</u>

In the previous example optimal α was found to be .266 with MSE = 19,447. An estimate of the standard deviation of the forecast errors is √9,447 = 139.5.

- If the forecast errors are approximately normally distributed about a mean of zero, there is about a *95%* chance that the actual observation will fall within 2 standard deviations of the forecast or within

$$\pm 2\sqrt{MSE} = \pm 2\sqrt{19{,}447} = \pm 2(139.5) = \pm 279$$

- For this example, the permissible absolute error is 279.
- If for any future forecast the magnitude of the error is greater than 279, α should be updated or a different forecasting method considered.
- The preceding discussion on tracking signals also applies to the smoothing methods yet to be discussed in the rest of the chapter.
- Simple exponential smoothing works well when the data vary about a level that changes infrequently.

Moving Averages and Smoothing Methods

- Whenever a sustained trend exists, exponential smoothing will lag behind the actual values over time.
- Holt's linear exponential smoothing technique, which is designed to handle data with a well-defined trend, addresses this problem and is introduced next.

2. Holt's Method

- This method is designed to handle time series data with a well defined trend.
- It is also called *Holt's linear exponential smoothing method*.
- When a trend in the time series is anticipated, an estimate of the current slope as well as the current level is required.
- Holt's technique smoothes the level and slope directly by using different smoothing constants for each.
- These smoothing constants provide estimates of level and slope that adapt over time as new observations become available.

Moving Averages and Smoothing Methods

ÖZYEĞİN UNIVERSITY

- One of the advantages of Holt's technique is that it provides a great deal of flexibility in selecting the rates at which the level and trend are tracked.
- The three equations used in Holt's method are:
  - The exponentially smoothed series, or current level estimate:

$$\boxed{L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})}$$

  - The trend estimate

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

  - Forecast p periods in the future

$$\boxed{\hat{Y}_{t+p} = L_t + pT_t}$$

L_t = new smoothed value (estimate of current level)
α = smoothing constant for the level ($0 < \alpha < 1$)
Y_t = new observation or actual value of series in period t
β = smoothing constant for trend estimate ($0 < \beta < 1$)

T_t = trend estimate
p = periods to be forecast into the future
$\hat{Y}_{t+p}$ = forecast for p periods into the future

Moving Averages and Smoothing Methods

- Usually, we start by setting the initial value for the level by $L_1 = Y_1$ and by setting the initial value for the trend by $T_1 = 0$.
- Another approach is to take the first five or six observations and set the initial value for the level by the mean of these observations and to set the initial trend value by the slope of the line that is fit to those five or six observations where the variable of interest Y is regressed on the time as its independent variable X.
- As with simple exponential smoothing, the smoothing constants α and β can be selected subjectively or by minimizing a measure of forecast error such as the *MSE*.
- Large weights result in more rapid changes in the component; small weights result in less rapid changes.
- Therefore, the larger the weights, the more the smoothed values follow the data; the smaller the weights, the smoother the pattern in the smoothed values.

Moving Averages and Smoothing Methods

- We could develop a grid of values of α and β and then select the combination that provides the lowest MSE.
- Most forecasting software packages use an optimization algorithm to minimize MSE.
- We might insist that $\alpha = \beta$, thus providing equal amounts of smoothing for the level and the trend.

Example: 9

In Example 6 simple exponential smoothing did not produce successful forecasts of Acme Tool Company saw sales. Because the data suggest that there might be a trend in these data, Holt's linear exponential smoothing is used to develop forecasts.

- To begin the computations shown in following table, two estimated initial values are needed, namely, the initial level and the initial trend value.
- The estimate of the level was set equal to the first observation.

Moving Averages and Smoothing Methods

- The trend was estimated to equal zero. The technique is demonstrated in the table for $\alpha = .3$ and $\beta = 1$

(1) Year	(2) t	(3) Y_t	(4) L_t	(5) T_t	(6) $\hat{Y}_{t+p}$	e_t	(1) Year	(2) t	(3) Y_t	(4) L_t	(5) T_t	(6) $\hat{Y}_{t+p}$	e_t
1996	1	500	500.0	0.0	500.0	0.0	1999	13	550	339.8	−3.3	249.7	300.3
	2	350	455.0	−4.5	500.0	−150.0		14	350	340.6	−2.9	336.5	13.5
	3	250	390.4	−10.5	450.5	−200.5		15	250	311.4	−5.5	337.7	−287.7
	4	400	385.9	−9.9	379.8	20.2		16	550	379.1	1.8	305.9	244.1
1997	5	450	398.2	−7.7	376.0	74.0	2000	17	550	431.7	6.9	381.0	169.0
	6	350	378.3	−8.9	390.5	−40.5		18	400	427.0	5.7	438.6	−38.6
	7	200	318.6	−14.0	369.4	−169.4		19	350	407.9	3.3	432.7	−82.7
	8	300	303.2	−14.1	304.6	−4.6		20	600	467.8	8.9	411.2	188.8
1998	9	350	307.4	−12.3	289.1	60.9	2001	21	750	558.7	17.1	476.8	273.2
	10	200	266.6	−15.2	295.1	−95.0		22	500	553.1	14.8	575.9	−75.9
	11	150	221.0	−18.2	251.4	−101.4		23	400	517.6	9.8	567.9	−167.9
	12	400	262.0	−12.3	202.8	197.2		24	650	564.2	13.5	527.4	122.6
							2002	25	850	—	—	577.7	273.3

$MSE = 20{,}515.5.$

- The computations leading to the forecast for period 3 are shown next.

  1. Update the exponentially smoothed series or level:

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})$$
$$L_2 = .3Y_2 + (1-.3)(L_{2-1} + T_{2-1})$$
$$L_2 = .3(350) + .7(500 + 0) = 455$$

Moving Averages and Smoothing Methods

2. **Update the trend estimate:**

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$
$$T_2 = .1(L_2 - L_{2-1}) + (1 - .1)T_{2-1}$$
$$T_2 = .1(455 - 500) + .9(0) = -4.5$$

3. **Forecast one period into the future:**

$$\hat{Y}_{t+p} = L_t + pT_t$$
$$\hat{Y}_{2+1} = L_2 + 1T_2 = 455 + 1(-4.5) = 450.5$$

4. **Determine the forecast error**

$$e_3 = Y_3 - \hat{Y}_3 = 250 - 450.5 = -200.5$$

- The forecast for period 25 is computed as follows:

1. **Update the exponentially smoothed series or level:**

$$L_{24} = .3Y_{24} + (1 - .3)(L_{24-1} + T_{24-1})$$
$$L_{24} = .3(650) + .7(517.6 + 9.8) = 564.2$$

2. Update the trend estimate:

$$T_{24} = .1(L_{24} - L_{24-1}) + (1 - .1)T_{24-1}$$
$$T_{24} = .1(564.2 - 517.6) + .9(9.8) = 13.5$$

3. Forecast one period into the future:

$$\hat{Y}_{24+1} = L_{24} + 1T_{24}$$
$$\hat{Y}_{25} = 564.2 + 1(13.5) = 577.7$$

- On the basis of minimizing the *MSE* over the period 1996 to 2002, Holt's linear smoothing (with $\alpha = .3$ and $\beta = 1$) does not reproduce the data any better than simple exponential smoothing that used α of .266.
- Forecasts for the actual sales for the first quarter of 2002 using Holt's smoothing and simple exponential smoothing are similar.
- A comparison of the *MAPEs* shows them to be about the same.
- To summarize:

$\alpha = .266$	$MSE = 19{,}447$	$MAPE = 32.2\%$
$\alpha = .3, \beta = .1$	$MSE = 20{,}516$	$MAPE = 35.4\%$

Moving Averages and Smoothing Methods

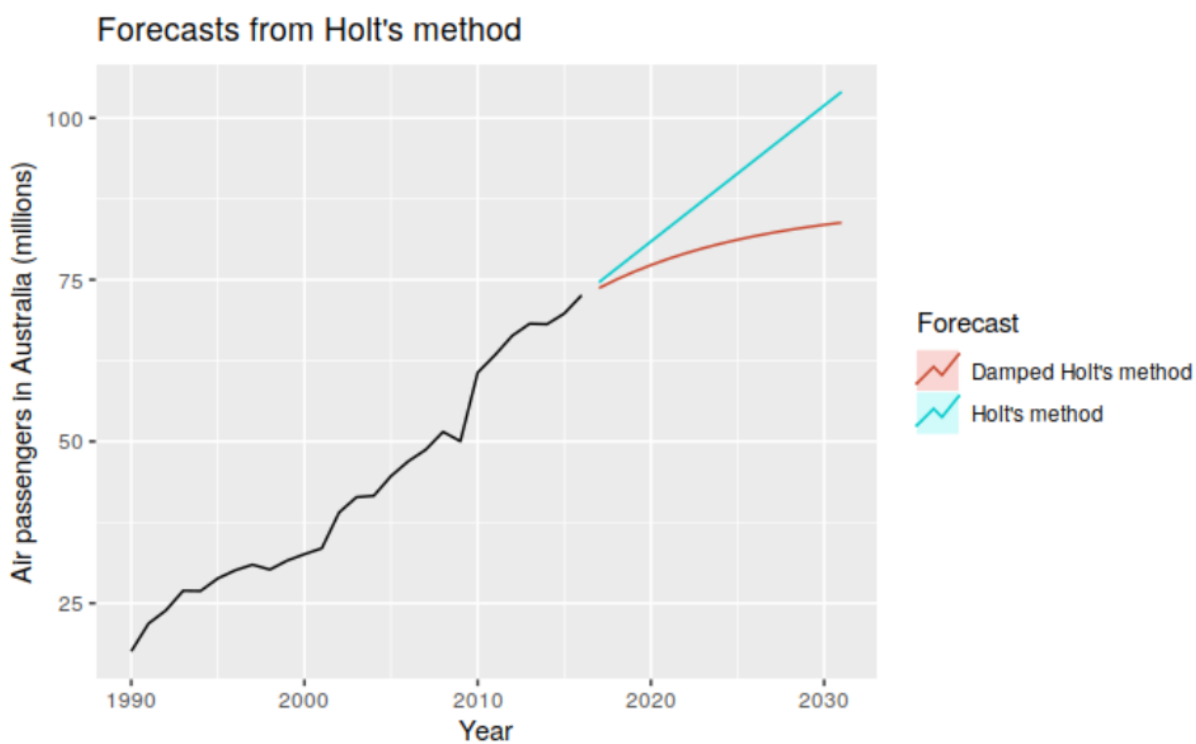
- Empirical evidence indicates that Holt's method tend to over-forecast, especially for longer forecast horizons.
- Holt's damped trend method

$$\hat{y}_{t+h|t} = \ell_t + (\phi + \phi^2 + \cdots + \phi^h) b_t$$
$$\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$$
$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}.$$

- Damping parameter $0 < \phi < 1$.
- If $\phi = 1$, identical to Holt's linear trend.
- As h → ∞, $\hat{y}_{T+h|T} \to \ell_T + \phi b_T / (1-\phi)$
- Short-run forecasts trended, long-run forecasts constant.
- In practice, ϕ is rarely less than 0.8 (otherwise damping has a very strong effect). $\phi = 1$ would imply a non-damped method. For these reasons, we usually restrict ϕ to a minimum of 0.8 and a maximum of 0.98.

Moving Averages and Smoothing Methods

- Forecasting total annual passengers of air carriers registered in Australia (millions of passengers, 1990–2016).
- For the damped trend method, $\phi = 0.9$
- We use (h=15) to highlight the difference between a damped trend and a linear trend. In practice, we would not want to forecast so many years ahead with only 27 years of data.



Moving Averages and Smoothing Methods

3. Winter's Method

- It is considered as an extension of Holt's method but it is designed to handle time series data with a well defined trend and seasonal variation.
- It is also called Winters' three parameter linear and seasonal exponential smoothing method.
- The four equations used in Winters' (multiplicative) smoothing are:
  - The exponentially smoothed series or level estimate:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

  - The trend estimate:

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

  - The seasonality estimate:

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma)S_{t-s}$$

Moving Averages and Smoothing Methods

- **Forecast p periods into the future:**

$$\hat{Y}_{t+p} = (L_t + pT_t)S_{t-s+p}$$

L_t = new smoothed value or current level estimate
α = smoothing constant for the level
Y_t = new observation or actual value in period t
β = smoothing constant for trend estimate
T_t = trend estimate
γ = smoothing constant for seasonality estimate
S_t = seasonal estimate
p = periods to be forecast into the future
s = length of seasonality
$\hat{Y}_{t+p}$ = forecast for p periods into the future

- **As with Holt's method, the weights α, β, and γ can be selected subjectively or by minimizing a measure of forecast error such as MSE.**
- **The most common approach for determining these values is to use an optimization algorithm to find the optimal smoothing constants.**

Example: 10

- Winters' technique is demonstrated in the following table for $\alpha = .4$, $\beta = .1$, and $\gamma = .3$ for the Acme Tool Company data. The value for α is similar to the one used for simple exponential smoothing in Example 6 and is used to smooth the data to create a level estimate. The smoothing constant β used to create a smoothed estimate of trend. The smoothing constant γ is used to create a smoothed estimate of the seasonal component in the data.

- The forecast for the first quarter of 2002 is 778.2.

- The computations leading to the forecast value for the first quarter of 2002, or period 25, are shown next.

Moving Averages and Smoothing Methods

Year	(1) t	(2) Y_t	(3) L_t	(4) T_t	(5) S_t	(6) $\hat{Y}_{t+p}$	(7) e_t
1996	1	500	415.459	−41.9541	1.26744	563.257	−63.257
	2	350	383.109	−40.9937	0.89040	328.859	21.141
	3	250	358.984	−39.3068	0.66431	222.565	27.435
	4	400	328.077	−38.4668	1.18766	375.344	24.656
1997	5	450	315.785	−35.8494	1.31471	367.063	82.937
	6	350	325.194	−31.3235	0.94617	249.255	100.745
	7	200	296.748	−31.0358	0.66721	195.221	4.779
	8	300	260.466	−31.5604	1.17690	315.576	−15.576
1998	9	350	243.831	−30.0679	1.35093	300.945	49.055
	10	200	212.809	−30.1632	0.94426	202.255	−2.255
	11	150	199.515	−28.4764	0.69259	121.863	28.137
	12	400	238.574	−21.7228	1.32682	201.294	198.706
1999	13	550	292.962	−14.1117	1.50886	292.949	257.051
	14	350	315.575	−10.4393	0.99371	263.306	86.694
	15	250	327.466	−8.2062	0.71385	211.335	38.665
	16	550	357.366	−4.3956	1.39048	423.599	126.401
2000	17	550	357.588	−3.9339	1.51763	532.584	17.416
	18	400	373.206	−1.9787	1.01713	351.428	48.572
	19	350	418.856	2.7843	0.75038	264.999	85.001
	20	600	425.586	3.1788	1.39628	586.284	13.716
2001	21	750	454.936	5.7959	1.55691	650.706	99.294
	22	500	473.070	7.0297	1.02907	468.626	31.374
	23	400	501.286	9.1484	0.76465	360.255	39.745
	24	650	492.469	7.3518	1.37336	712.712	−62.712
2002	25	850	—	—	—	778.179	—
	26	600	—	—	—	521.917	—
	27	450	—	—	—	393.430	—
	28	700	—	—	—	716.726	—

MSE = 7,636.86.

1. **The exponentially smoothed series or level estimate**

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$L_{24} = .4 \frac{Y_{24}}{S_{24-4}} + (1 - .4)(L_{24-1} + T_{24-1})$$

$$L_{24} = .4 \frac{650}{1.39628} + (1 - .4)(501.286 + 9.148)$$

$$L_{24} = .4(465.52) + .6(510.434) = 492.469$$

Moving Averages and Smoothing Methods

2. **The trend estimate:**

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$
$$T_{24} = .1(L_{24} - L_{24-1}) + (1 - .1)T_{24-1}$$
$$T_{24} = .1(492.469 - 501.286) + .9(9.148)$$
$$T_{24} = .1(-8.817) + .9(9.148) = 7.352$$

3. **Seasonality estimate:**

$$S_t = \gamma\frac{Y_t}{L_t} + (1-\gamma)S_{t-s}$$
$$S_{24} = .3\frac{Y_{24}}{L_{24}} + (1 - .3)S_{24-4}$$
$$S_{24} = .3\frac{650}{492.469} + .7(1.39628)$$
$$S_{24} = .3(1.3199) + .9774 = 1.3734$$

4. **Forecast $p = 1$ period into the future:**

$$\hat{Y}_{24+1} = (L_{24} + 1T_{24})S_{24-4+1}$$
$$\hat{Y}_{25} = (492.469 + 1(7.352))1.5569 = 778.17$$

Moving Averages and Smoothing Methods

- For the parameter values considered, the Winters' technique is better than both of the previous smoothing procedures in terms of minimizing *MSE*.
- When the forecasts for the actual sales for the first quarter of 2002 are compared, the Winters' technique also appears to do a better job.
- Winters' exponential smoothing method seems to provide adequate forecasts for the Acme Tool Company data.
- Winters' method provides an easy way to account for seasonality when data have a seasonal pattern.
- An alternative method consists of first deseasonalizing or seasonally adjusting the data.
- Deseasonalizing is a process that removes the effects of seasonality from the raw data .
- Exponential smoothing is a popular technique for short-run forecasting. Its major advantages are low cost and simplicity.

Moving Averages and Smoothing Methods

- Holt-Winters additive method: When the seasonal variations are roughly constant through the series

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$$
$$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$$
$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$
$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m},$$

- k = integer part of $(h - 1)/m$. Ensures estimates from the final year are used for forecasting.
- Parameters: $0 \le \alpha \le 1, 0 \le \beta^* \le 1, 0 \le \gamma \le 1 - \alpha$ and

m = period of seasonality (e.g. m = 4 for quarterly data).

Moving Averages and Smoothing Methods

- Seasonal component is usually expressed as:

$$s_t = \gamma^*(y_t - \ell_t) + (1 - \gamma^*)s_{t-m}.$$

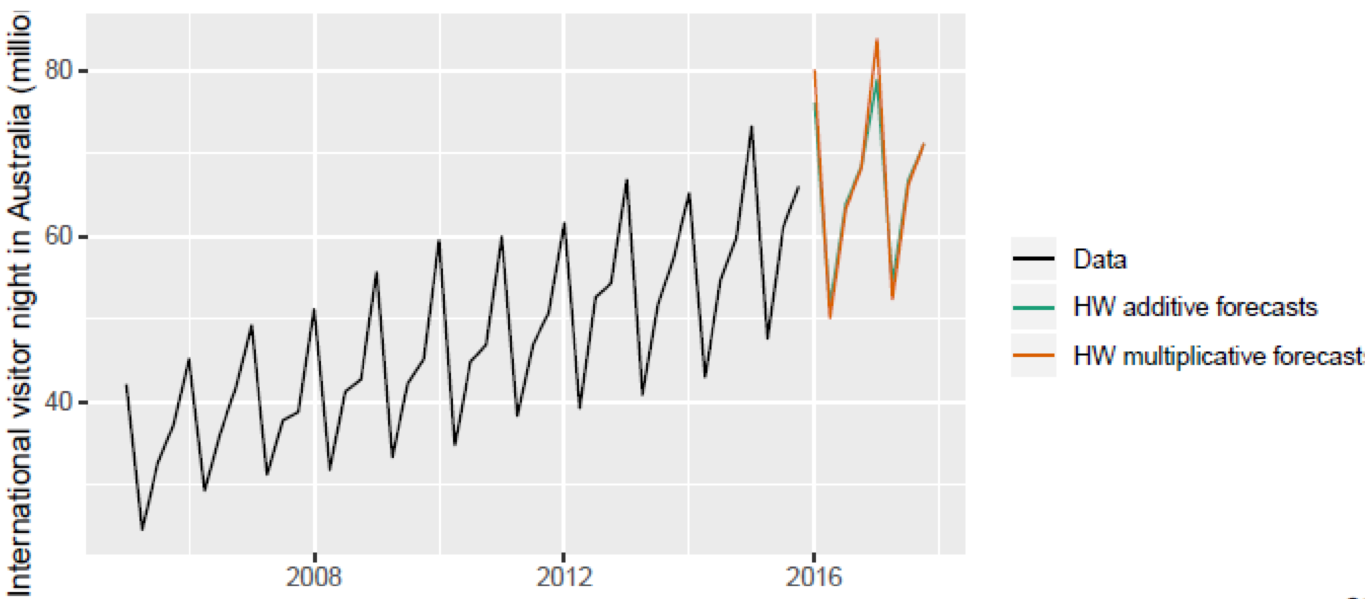
- Substitute in for ℓ_t:

$$s_t = \gamma^*(1-\alpha)(y_t - \ell_{t-1} - b_{t-1}) + [1 - \gamma^*(1-\alpha)]s_{t-m}$$

- We set $\gamma = \gamma^*(1 - \alpha)$
- The usual parameter restriction is $0 \leq \gamma^* \leq 1$, which translates to $0 \leq \gamma \leq (1 - \alpha)$.

Moving Averages and Smoothing Methods

- Quarterly visitor nights in Australia spent by international tourists (peaks observed in the March quarter of each year)



- Additive model: $\alpha = 0.306, \beta^* = 0.0003, \gamma = 0.426, RMSE = 1.763$
- Multiplicative model: $\alpha = 0.441, \beta^* = 0.030, \gamma = 0.002, RMSE = 1.576$
- Can we compare the RMSE from both models?

Moving Averages and Smoothing Methods

- Holt-Winters damped method
- Often the single most accurate forecasting method for seasonal data:

$$\hat{y}_{t+h|t} = [\ell_t + (\phi + \phi^2 + \cdots + \phi^h)b_t]s_{t+h-m(k+1)}$$

$$\ell_t = \alpha(y_t/s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$$

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$$

$$s_t = \gamma \frac{y_t}{(\ell_{t-1} + \phi b_{t-1})} + (1-\gamma)s_{t-m}$$

Moving Averages and Smoothing Methods

- Taxonomy of exponential smoothing methods

Trend Component		Seasonal Component N (None)	A (Additive)	M (Multiplicative)
N	(None)	(N,N)	(N,A)	(N,M)
A	(Additive)	(A,N)	(A,A)	(A,M)
A_d	(Additive damped)	(A_d,N)	(A_d,A)	(A_d,M)

  - (N,N): Simple exponential smoothing
  - (A,N): Holt's linear method
  - (Ad,N): Additive damped trend method
  - (A,A): Additive Holt-Winters' method
  - (A,M): Multiplicative Holt-Winters' method
  - (Ad,M): Damped multiplicative Holt-Winters' method