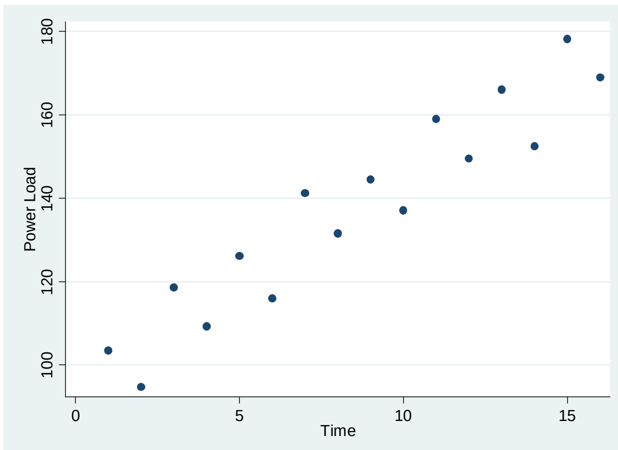
MA397 – Smoothing Methods for Forecasting in Time Series

Moving Average Smoother

For this exercise we will be using the *qtrpower.dta* dataset found on the course webpage at http://www.colby.edu/personal/l/lobrien/ma397.html.

The data consist of quarterly power demands over a four-year period. "Power" is the lower load, "year" is the year of measurement," and "time" is the quarter of measurement. A plot of "power" against "time" is below:



Notice that there is a clear secular trend, as well as a possible seasonal variation. In order to apply the smoothing techniques that we have learned in class, we need to first set this up as a time series data set. To do this, click on **Statistics > Time series > Setup and utilities > Declare data set to be time series data**. Enter "time" in the time variable box (note that the values of this variable have to be unique in each observation), and click on "quarterly" for the format. This is necessary when performing seasonal adjustments.

```
. tsset time, quarterly
        time variable:  time, 1960q2 to 1964q1
                delta:  1 quarter
```

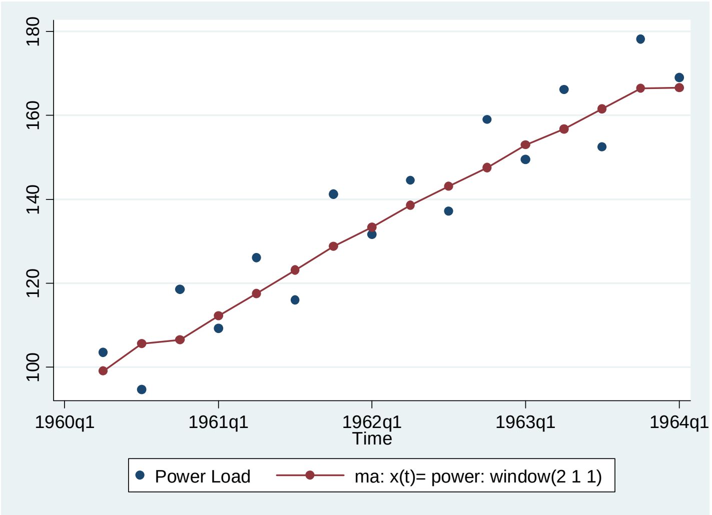
We will first apply a 4-point moving average smoother to these data. We use a 4-point moving average since the data are obtained quarterly. Thus, for time 3, we obtain an average of the first 4 power load observations (we use the previous two, current, and first future measurements). To do this in Stata, click on **Statistics > Time series > Smoothers/univariate forecasters > Moving average filter**. We will need to generate a new variable to hold the moving averages. Enter "mapower" in the new variable box, and enter "power" in the expression to smooth box. Enter 2 lagged terms, click on "current observation" and include the current observation, and click on "number of lead terms" and enter 1.

```
. tssmooth ma mapower = power, window(2 1 1)
The smoother applied was
     (1/4)*[x(t-2) + x(t-1) + 1*x(t) + x(t+1)]; x(t)= power
```

We can now list the data and see what was done:

```
. list

     +---------------------------------+
     | year     time   power   mapower |
     |---------------------------------|
  1. | 1999   1960q2   103.5      99.1 |
  2. | 1999   1960q3    94.7     105.6 |
  3. | 1999   1960q4   118.6   106.525 |
  4. | 1999   1961q1   109.3   112.175 |
  5. | 2000   1961q2   126.1     117.5 |
     |---------------------------------|
  6. | 2000   1961q3     116    123.15 |
  7. | 2000   1961q4   141.2   128.725 |
  8. | 2000   1962q1   131.6   133.325 |
  9. | 2001   1962q2   144.5     138.6 |
 10. | 2001   1962q3   137.1    143.05 |
     |---------------------------------|
 11. | 2001   1962q4     159   147.525 |
 12. | 2001   1963q1   149.5   152.925 |
 13. | 2002   1963q2   166.1   156.775 |
 14. | 2002   1963q3   152.5   161.575 |
 15. | 2002   1963q4   178.2    166.45 |
     |---------------------------------|
 16. | 2002   1964q1     169  166.5667 |
     +---------------------------------+
```

Notice that the first two observations are calculated as truncated moving averages in Stata. Also, the time variable has been automatically converted to a year/quarter format with 0 being equivalent to the first quarter 1960. The actual value is of little relevance here and we could convert it if necessary by changing the start value of "time." We can plot the new smoothed data with the original data:



Note that the smoothed data show the secular trend, but have smoothed out the seasonal variation. If we wanted to forecast the value for the next quarter (beyond the end of the data) we would have to do so visually.

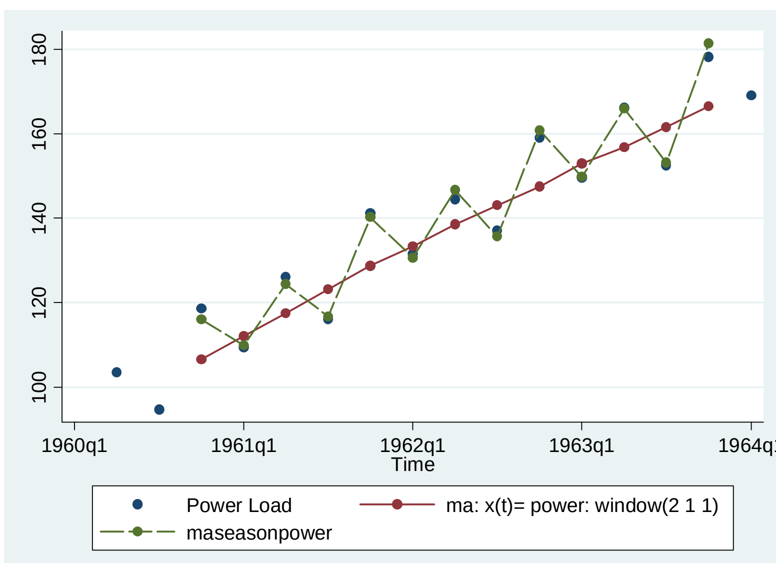
Recall that we can also account for the seasonal variation when using the moving average smoother. However, Stata does not allow this to be done through the moving average smoother. Instead, we can calculate the seasonal average through the application of a few quick commands. First generate a variable that tells which quarter (1, 2, 3, or 4) the observation is from. We should also set the first two smoothed values, and the last smoothed values, to be missing. Then we can calculate the ratio of the observed value to the moving average and get the average ratio by quarter:

```
. replace mapower = . if time==1 | time==2 | time==_N
(3 real changes made, 3 to missing)

. gen ratio=power/mapower
(3 missing values generated)

. by quarter: egen seasonalindex=mean(ratio)
```

Note that the seasonal index is a percentage and should be multiplied by 100 to reflect this. However, to obtain the seasonally adjusted moving average we will use the current fractional values:

```
. gen maseasonpower=mapower*seasonalindex
(3 missing values generated)
```

Let's plot the two moving average smoother over the original data:



We can make forecasts for future values of time by visually analyzing this plot.

Exponential Smoothing

The advantage of exponential smoothing over the moving average smoother is that future observations are not used in the smoothing window. Thus we do not lose observations on the right side of the time series (we lost one for the lead term in the 4-point moving average earlier).

To apply an exponential smoother in Stata, click on **Statistics > Time series > Smoothers/univariate forecasters > Single exponential smoothing**. Enter the new variable name that will hold the exponentially smoothed values and the name of the variable to smooth in the "expression to smooth" box. Click on the "specify smoothing parameter" box and enter 0.7 (for this example).

```
. tssmooth exponential exppower = power, parms(.7)

exponential coefficient  =       0.7000
sum-of-squared residuals =       3084.2
root mean squared error  =       13.884
```

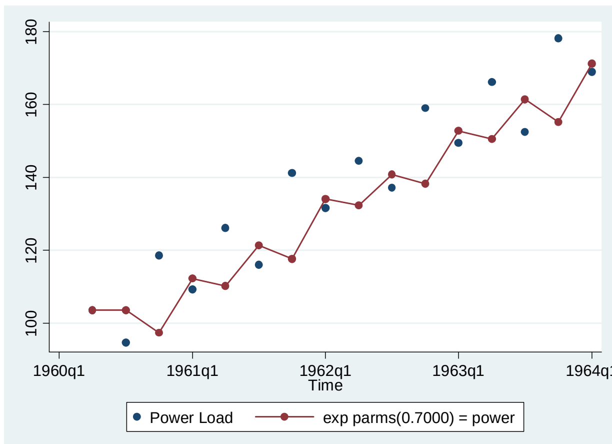
If you want Stata to choose the smoothing constant to minimize the sum of squared forecast errors, you may let it default. In this case it defaults to w = 0.6024.

Again, note that in the first observation, Stata has entered a different value from that observed. By default, it takes the average of the first half of the sample as its firth smoothed value. If you want to specify the starting value, you may enter the option "s0(xxx)" where xxx is the starting value. Do this for the starting value of 166.1 (note that the "optimal" smoothing constant is 0.6098 in this case). This may also be done by clicking on "Initial value for the recursion" in the dialog box and entering the value there. You may also click the "Periods for out of sample forecast" box and enter the number of periods you wish to forecast.

```
. tssmooth exponential exppower2 = power, s0(103.5) parms(.7) forecast(6)

exponential coefficient  =       0.7000
sum-of-squared residuals =         2846
root mean squared error  =       13.337

. list year time power quarter exppower exppower2

     +------------------------------------------------------+
     | year     time   power   quarter   exppower   exppow~2 |
     |------------------------------------------------------|
  1. | 1999   1960q2   103.5         1    117.625      103.5 |
  2. | 1999   1960q3    94.7         2   107.7375      103.5 |
  3. | 1999   1960q4   118.6         3   98.61125      97.34 |
  4. | 1999   1961q1   109.3         4   112.6034    112.222 |
  5. | 2000   1961q2   126.1         1    110.291   110.1766 |
     |------------------------------------------------------|
  6. | 2000   1961q3     116         2   121.3573    121.323 |
  7. | 2000   1961q4   141.2         3   117.6072   117.5969 |
  8. | 2000   1962q1   131.6         4   134.1222   134.1191 |
  9. | 2001   1962q2   144.5         1   132.3567   132.3557 |
 10. | 2001   1962q3   137.1         2    140.857   140.8567 |
     |------------------------------------------------------|
 11. | 2001   1962q4     159         3   138.2271    138.227 |
 12. | 2001   1963q1   149.5         4   152.7681   152.7681 |
 13. | 2002   1963q2   166.1         1   150.4804   150.4804 |
```

```
 14. | 2002   1963q3   152.5         2   161.4141    161.4141 |
 15. | 2002   1963q4   178.2         3   155.1742    155.1742 |
     |---------------------------------------------------------|
 16. | 2002   1964q1     169         4   171.2923    171.2923 |
 17. |    .   1964q2       .         .          .    169.6877 |
 18. |    .   1964q3       .         .          .    169.6877 |
 19. |    .   1964q4       .         .          .    169.6877 |
 20. |    .   1965q1       .         .          .    169.6877 |
     |---------------------------------------------------------|
 21. |    .   1965q2       .         .          .    169.6877 |
 22. |    .   1965q3       .         .          .    169.6877 |
     +---------------------------------------------------------+
```

Let’s plot the exponentially smoothed data using the 103.5 start value (exppower2) against the original data:

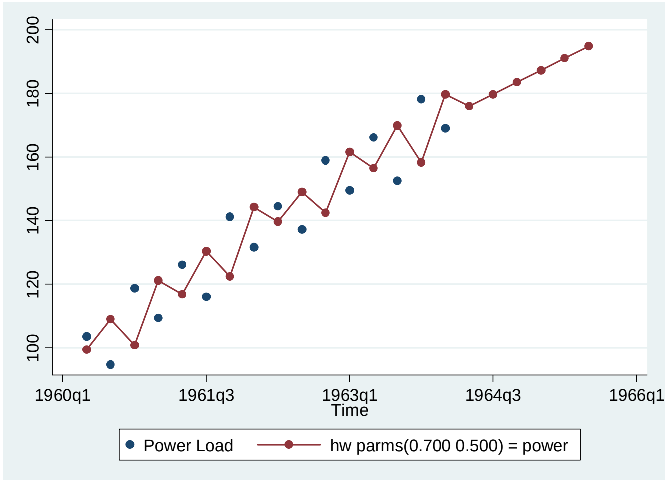


The forecast for any future time period is 171.3.

Holt-Winters Exponential Smoothing

The Holt-Winters model improves upon the exponential smoothing by allowing for forecast adjustments due to secular trend and seasonal variation. If we want Stata to perform this smoothing not accounting for seasonal variation, click on **Statistics > Time series > Smoothers/univariate forecasters > Holt-Winters nonseasonal smoothing.** You may enter values for the smoothing constants (0.7 and 0.5 in this case) or let them default to minimize the sum of squared forecast errors. Enter the name of the variable to hold the smoothed values (hwpower), and enter “power” in the expression to smooth box. You may enter initial values, although Stata will treat them as values for time 1, so that the results will not agree with those in the textbook. You can also enter the number of periods to forecast (enter 6 here).

```
. tssmooth hwinters hwpower = power, parms(.7 .5) forecast(6)

Specified weights:
                   alpha = 0.7000
                    beta = 0.5000
sum-of-squared residuals = 2984.343
 root mean squared error = 13.65728
```



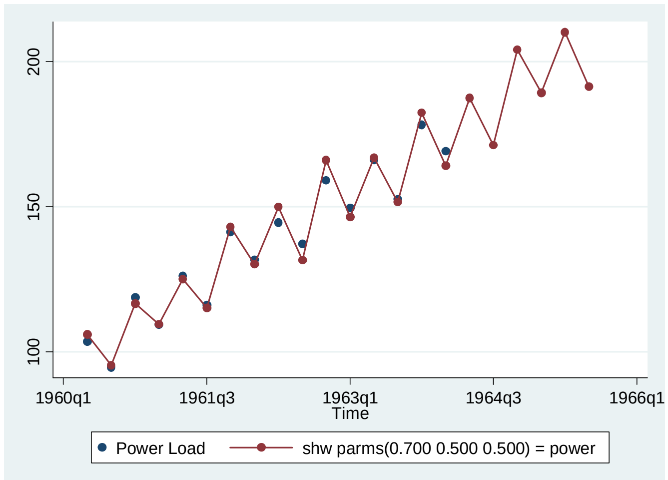
We can add a seasonal adjustment by choosing **Statistics > Time series > Smoothers/univariate forecasters > Holt-Winters seasonal smoothing.** The options are similar (the new variable should be "hwseasonpower") except that you must enter a smoothing constant for the seasonal component (0.5 in this case). We also need to enter the periodicity of the seasonal effect (4 in this case) and can enter the number of post-sample forecast periods (enter 6).

```
. tssmooth shwinters hwseasonpower = power, parms(.7 .5 .5) forecast(6) period(4)

Specified weights:
                   alpha = 0.7000
                    beta = 0.5000
                   gamma = 0.5000
sum-of-squared residuals = 180.9611
 root mean squared error = 3.363045
```

```
. list year time power quarter hwpower  hwseasonpower

     +--------------------------------------------------------+
     | year     time   power   quarter    hwpower   hwseas~r |
     |--------------------------------------------------------|
  1. | 1999   1960q2   103.5         1   99.36667   106.0126 |
  2. | 1999   1960q3    94.7         2   108.9233    95.2976 |
  3. | 1999   1960q4   118.6         3   100.6522   116.5356 |
  4. | 1999   1961q1   109.3         4   121.1826   109.4087 |
  5. | 2000   1961q2   126.1         1   116.6728   125.0791 |
     |--------------------------------------------------------|
  6. | 2000   1961q3     116         2   130.3794   115.0656 |
  7. | 2000   1961q4   141.2         3   122.3886    143.007 |
  8. | 2000   1962q1   131.6         4   144.2153    130.175 |
  9. | 2001   1962q2   144.5         1    139.628    149.865 |
 10. | 2001   1962q3   137.1         2    148.987   131.4745 |
     |--------------------------------------------------------|
 11. | 2001   1962q4     159         3   142.4543   166.0786 |
 12. | 2001   1963q1   149.5         4   161.6154   146.4491 |
 13. | 2002   1963q2   166.1         1   156.4734   166.9071 |
 14. | 2002   1963q3   152.5         2   169.9201   151.5628 |
 15. | 2002   1963q4   178.2         3   158.3371   182.3716 |
     |--------------------------------------------------------|
 16. | 2002   1964q1     169         4   179.8042    164.098 |
 17. |    .   1964q2       .         .   176.0229    187.369 |
 18. |    .   1964q3       .         .   179.8045   171.1787 |
 19. |    .   1964q4       .         .   183.5861   203.9775 |
```

```
 20. |     .   1965q1       .          .   187.3677   189.2283 |
     |-------------------------------------------------------|
 21. |     .   1965q2       .          .   191.1493   210.0077 |
 22. |     .   1965q3       .          .   194.9309   191.2547 |
     +-------------------------------------------------------+
```

The plot of the seasonally adjusted Holt-Winters smoothed values is below:



Stata will calculate the root mean squared forecast error when performed out of sample forecasts but does not calculate the mean absolute deviation or mean absolute percent error by default. You can calculate them if you wish by using the “generate” command if you have the future data.