

Week 2 February 1st, 2006
RATS Lecture 2 – ARMA and ARIMA Modeling
This accompanies Lecture 2.

Open the USDATA data set.

We'll use monthly presidential approval – the monthly percentage of people who say they approve of the job _____ is doing as president.

Do we have any theoretical reasons for thinking this series is stationary or non-stationary?

Take a look at the values.

print / pres

To diagnose whether a series is stationary or not, we begin by looking at a graph of the series.

```
graph(key=loright,header='Presidential Approval',  
subhead='1978-1997') 1  
# pres
```

Diagnosing autoregressive and moving average parameters.

Practice with presidential approval and then nret – national retrospections.

Any *a priori* theoretical ideas about what the model could be?

See presapproval data set.

Check initial ACF & PACF – indicate a unit-root.

Running a simple (1,0,0) model indicates again that the series should be differenced.

Run diagnostic tests on that – what is the AIC, SBC compared with rival models?

Take a close look at the program:

```
calendar 1978 1 12  
allocate 1997:12  
open data usdata.txt  
data(format=free,org=obs) / pres npros3 npros1 npros2 nret ppros pret dstorm $  
iranp rally19 iranc reagan tien bush irann clinton hinkley  
  
source bjidnt.src (here we read in whatever source files we will need)  
  
statistics pres
```

```
graph(key=loright,header='Presidential Approval',subhead='1978-1997') 1
# pres          (a graph of the series – helps decide about stationarity)

@bjident pres  (“@” invokes a procedure, then name the procedure and the series)
                There are options here we could use also.
```

This next part does essentially what the bjident procedure does – but now we can see the steps:

```
correlate(number=50,partial=prespart,stderrs=auto,$
qstats) pres / presaut    I tell RATS how many lags, name the partial autocorrelations and the S.E.s.
                           presaut is the name of the series of autocorrelations I made.

print / auto              Look at standard errors.

graph(style=bargraph,key=loright,nodates,number=0,$
max=1.0,min=-1.0) 3
# presaut
# prespart
# auto

boxjenk(ar=1) pres / resids    Run a noise model to estimate parameters.
compute aic = 239*log(4756.23) + 2*1    Akaike Information Criterion.
compute sbc = 239*log(4756.23) + 1*log(239)    Schwartz Bayesian Criterion.
compute rms = (sqrt(4756.23))/239    Residual Mean Square.
```

```
display 'aic = ' aic ' sbc = ' sbc ' rms = ' rms    Display diagnostics.
```

```
diff pres / presd    Now try with a differenced series.
correlate(number=25,partial=presdpart,stderrs=auto2,$
qstats) presd / presdaut
print / auto2
graph(style=bargraph,key=loright,nodates,number=0,$
max=1.0,min=-1.0) 3
# presdaut
# presdpart
# auto2
```

```
@bjident presd
```

Now try for National Retrospections

```
correlate(number=50,partial=nretpart,stderrs=auto3,$
qstats) nret / nretaut
```

Q Stats will be helpful here: $Q = T(T + 2) \sum_{k=1}^s r_k^2 / (T - k)$ which is $\sim \chi^2$.

```
print / auto3
graph(style=bargraph,key=loright,nodates,number=0,$
max=1.0,min=-1.0) 3
# nretaut
# nretpart
# auto3
```

```
@bjident nret
```

```
diff nret / nretd
```

```
@bjident nretd
```

Looking at the differenced series, there are significant spikes in the PACF.

**In the next line note the "||2,5||". This tells RATS I want to estimate MA parameters
** at the 2nd and 5th lag. Without the lines, RATS will estimate at lags 1,2,3, etc...
** So, to get different lags, use "||lag||". The "|" is shift--back-slash.

```
boxjenk(diffs=1,ar=1,ma=||2,5||) nret / resid
```

```
compute aic = 235*log(19094.41) + 2*1  
compute sbc = 235*log(19094.41) + 1*log(235)  
compute rms = (sqrt(19094.41))/235
```

```
display 'aic = ' aic ' sbc = ' sbc ' rms = ' rms
```

The best model looks to be a $(1,1,2^{2.5})$.

Forecasting:

Making forecasts is easy:

Backcasting:

** Forecasts -- here we back-cast the last four observations of the NRET series

```
boxjenk(diffs=1,ar=1,ma=||2,5||,define=eq1) nret 5 236 nretres  
the "define=eq1" is to give a name to the equation we are using to forecast.  
then I state the series, the first date of the series (5 because of the MA5, and the  
last date of the series. Last comes a name I give for the vector of residuals.
```

```
forecast(print) 1 4 237
```

I have to tell RATS to print out the forecast.

1 means compute the forecasts for the 1st equation in memory.

4 is the number of forecasts I am asking for.

237 means to begin the forecasts with month 237.

```
# eq1
```

do equation 1.

** Compare with actual values

```
print 97:09 97:12 nret
```

Forecasting:

** Forecasts -- Forward casting

```
boxjenk(diffs=1,ar=1,ma=||2,5||,define=eq2) nret 5 240 nretres2
```

```
forecast(print) 1 4 241
```

```
# eq2
```

** No values in this data set with which to compare these forecasts.