Exercises - Day I

Ashish Rajbhandari
Stata
arajbhandari@stata.com

February 08 & 09, 2016

1 Exercise 1

1. We have data on the U.S. unemployment rate in `unrate.data` obtained from the FRED database at the St. Louis FED. The two variables in this dataset are `date` and `UNRATE`.

   - `date` is a string variable. Create a new variable called `td` that records the `date` in a Stata daily format.
   - Format `td` as a daily date variable and list out the first 10 observations on `date` and `td` to verify the previous step.
   - Create a new variable called `tm` that records the `date` in a Stata monthly format.
   - Format `tm` as a monthly date variable and list out the first 10 observations on `date`, `td`, and `tm` to verify the previous step.
   - Set `tm` as your time variable.
   - Generate a new variable called `dunrate` containing the first difference of `UNRATE`.
   - Use `tsline` to draw a line graph of `dunrate`.

2. The dataset `returns.dta` has data on returns from a market that is always closed on Sundays and has some irregular closures during the first 200 days of 2011, which is our sample period.

   We want to construct a business calendar that handles the missing observations on Sundays as well as for those on irregular closures.

   - Load the `returns` dataset, use the `describe` command to describe the data. `td` is a Stata date variable. Use `codebook` to get more information on the date variable
   - Write a `.stbcal` file called returns.stbcal with a header information, date range that also handles the missing Sundays.

- Use `tsfill` command to fill in the `td` for all missing observations.

We want to list out `td` for days that are not Sundays but for which returns contain missing values.

- Generate a day of the week variable `day` using the `dow()` function on `td`
- Use the `list` command and the `missing()` function to list out `td` for these days.
- We want a clean list without observation numbers, so that we can just copy the dates from the output and paste them into our `.stbal` file. Add the options clean and noobs to the list command from above to get the clean list without observation numbers.
- Use list of omitted dates to complete `returns.stbcal` and save.
- Create a business calendar date `tb` from `td` using the function `bofd()`. Verify everything is correct and `tsset` the new date variable.

2 Exercise 2

1. In this example we use simulated data to illustrate some properties of multiplicative ARMA model.

   - Read in the `sarima.dta` dataset, describe the data, and make sure that it is properly `tsset`. The frequency is not given in the generic time variable, but it is quarterly data. Make sure to format the time variable appropriately.

     Estimate the parameters of the multiplicative ARMA model

     $$(1 - \phi_1 L - \phi_2 L^2)(1 - \phi_{s4} L^4) y_t = (1 + \theta_1 L)(1 + \theta_{s4} L^4)\epsilon_t$$

   - Using standard algebra, write out the underlying ARMA model corresponding to this multiplicative ARMA model. (Do not plug in the estimates, do the algebra in terms of the parameters ϕ_1, ϕ_2, ϕ_{s4}, θ_1, and θ_{s4}).
   - Use the estimates of the multiplicative ARMA parameters to calculate what the estimated ARMA parameters should be.
   - Estimate the parameters of the underlying ARMA model. Are the estimates close to the ones you computed?

2. The dataset `paynsa.dta` contains data on the number of employed people in the U.S. The data has not been deseasonalized.

   - Load and describe the data. Notice that `date` is a string date. Convert the string date into Stata date with the appropriate frequency.

- Graph the log of employment contained in `ln_paynsa`. Is there a trend?
- Graph the log differenced series contained in `dln_paynsa`. Is there a trend?
- Which of the following models produce the smallest BIC?
  (a) $(1 - \phi_1 L - \phi_{12} L^{12}) y_t = (1 + \theta_1 L + \theta_{12} L^{12}) \epsilon_t$
  (b) $(1 - \phi_1 L)(1 - \phi_{s12} L^{12}) y_t = (1 + \theta_1 L) \epsilon_t$
  (c) $(1 - \phi_1 L)(1 - \phi_{s12} L^{12}) y_t = (1 + \theta_1 L)(1 + \theta_{s12} L^{12}) \epsilon_t$
  where $y_t =$ `dln_paynsa`.
- Do you trust the results? (Hint: Does any of the parameter estimates in your selected model look suspicious?)
- Reestimate the model adjusting for seasonlity.