Demographics Problem Set

TPRU, Evans and DeBacker
Thursday, Aug. 22

1. **Indian demographics.** In the `OG-India` model repository, there is an updated module called `demographics.py`.

   - Navigate to your local `OG-India` directory on your computer that is a clone of your fork of the main repository.
   - Make sure that you have the most updated version of the master branch of the main repository merged into the master branch of your local and remote fork of the repository. Make sure you are in the master branch of your local repository, and type the following commands.
     - `git fetch upstream` (this collects any changes from the main repository master branch)
     - `git merge upstream/master` (this merges those changes into your local [on your laptop] master branch of your fork)
     - `git push origin master` (this pushes those changes into your remote [in the cloud] mast branch of your fork)
   - Create a new branch in which you will make the following changes by typing `git checkout -b demog`
   - Activate the `ogindia-dev` environment using the command `source activate ogindia-dev` if you are using a Mac OS X machine or using the command `activate ogindia-dev` if you are on a Windows machine.

   Use the functions in the module `demographics.py` in combination with additional code that you write to do the following exercises. You can import the demographics module by typing the following command.

   ```
   from ogindia-dev import demographics as demog
   ```

   You may want to open a Python interactive development environment (IDE) such as iPython, Spyder, or a Jupyter notebook to do some of these exercises. I recommend a Jupyter notebook.

   (a) Look through the `get_fert` function in the `demographics.py` module. Notice that it has been updated with the Indian fertility rate data. Use the `get_fert()` function with inputs `totpers=100`, `min_yr=1`, `max_yr=100`, and `graph=True` to create a NumPy array that you name `fert_rates`. Report the shape of the `fert_rates` array its the mean value. Show your plot of the fertility rates by age.

(b) Look through the `get_mort` function in the `demographics.py` module. Notice that it has been updated with the Indian mortality rate data. Use the `get_mort()` function with inputs `totpers=100`, `min_yr=1`, `max_yr=100`, and `graph=True` to create a NumPy array that you name `mort_rates`. Report the shape of the `mort_rates` array its the mean value. Show your plot of the mortality rates by age.

(c) We want to estimate the immigration rates by age i_s, but we do not have data on those rates. So we will have to use the population law of motion equations to estimate the immigration rates. However, the Indian population data has some strange properties. Use the read in the data file india_pop_data.csv that has population totals for India by age for the two years 2001 and 2011. Plot these two series by age—one line plot for 2001 population by age and one line plot for 2011 population by age—on the same axes. Make sure to include a legend and axis labels in your plot.

(d) Use a 5th order polynomial regression of the following form to smooth the 2001 population data by age.

$$pop_{2001,s} = \beta_0 + \beta_1 age_s + \beta_2 age_s^2 + \beta_3 age_s^3 + \beta_4 age_s^4 + \beta_5 age_s^5 + \varepsilon_s$$

Plot the 2001 population data along with the predicted population values from your regression with population on the y-axis and age on the x-axis.

(e) Do the same operations from part (d) on the 2011 population data. Use a 5th order polynomial regression of the following form to smooth the 2011 population data by age.

$$pop_{2011,s} = \gamma_0 + \gamma_1 age_s + \gamma_2 age_s^2 + \gamma_3 age_s^3 + \gamma_4 age_s^4 + \gamma_5 age_s^5 + \mu_s$$

Plot the 2011 population data along with the predicted population values from your regression with population on the y-axis and age on the x-axis.

(f) Plot your two predicted population series for 2001 and 2011 the predicted population values from your regression with population on the y-axis and age on the x-axis.

(g) Add the appropriate lines of code to the `demographics.py` module to the `get_imm_resid()` function to estimate the smoothed population functions you computed in parts (d), (e), and (f) [HINT: take `pop_2001` and `pop_2011` and from those estimate something like `pop_2001_smth` and `pop_2001_smth`.]. Run the `get_pop_objs()` function using the following command.

```
pop_objs = demog.get_pop_objs(20, 80, 320, 1, 100, 2020)
```

Make a plot of the population distribution at time periods $t = 2011, 2020, 20, 80, 320, 1, 100, 2020, 2060, 2100, SS$. Plot the path of the population growth rate from $t = 2020$ to $t = 2100$. Plot the immigration rates by age (This object is a matrix. Just plot the rates for $t = 2020$.)