Assignment for the EDSD 2022 course "Population projections"

Ugofilippo Basellini[1]

[1]Max Planck Institute for Demographic Research (MPIDR)

June 8, 2022

Please send your solutions at basellini@demogr.mpg.de with a single `.pdf` file containing your answers, as well as the `R` codes and data to generate the results. Ideally you can send an `R Markdown` document.

Reminder: choose only 5 of the following exercises, and work in team with your group members (Group 5 will need to hand in only 3 exercises).

Groups

- **Group 1**: Daniela, Augusto and Patrik
- **Group 2**: Dorothee, Sebastian and Daniele
- **Group 3**: Teodora, Elizabeth and Chiara
- **Group 4**: Agnese, Donata and Saskia
- **Group 5**: Camila, Nadia and Mariana
- **Group 6**: Rolf, Dorean, Abigail

Question 1

Go back to the first exercise that we saw today (constant exponential growth model), and compute $\hat{t}$ (time when population b surpasses a) without using the approximation formula, i.e. use $N(T) = N(0)e^{rT}$ instead of $N(T) = N(0)(1 + r)^T$. Is $\hat{t}$ higher or lower? When does the approximation formula works better?

Question 2

Take the population time-series for any country (data from HMD or UN), and answer to the following questions:

1. Estimate the population growth rate using two different time intervals (e.g. the last 10 and the last 15 years).

2. Project the population 30-year ahead using the constant exponential model and the two different growth rates
3. Make an assumption about one of the demographic components, and compare this projection with the baseline one
4. Assume that your observed data ends 15 years before the last available data point. Project the population 15-year ahead with the two growth rates estimated in point 1. Which of the two projections is more accurate?

Question 3

Take the population for any country (for example, use data from the HMD), as well as data for person-years and fertility rates (for example, take data from HMD and HFD). Specifically, take these data 5 years before the last available date (e.g. if the last available year is 2020, take the data in 2015). Divide the population into 5 years age groups and project the population 5-year ahead for each sex separately. Plot and compare your results with the observed population in the last available data (e.g. in 2020). Are your projections close to the observed population? At what ages do you see larger differences? Briefly provide a motivation for this discrepancy.

Question 4

Starting from the following equation

$$N_0^F(t+5) = B^F[t, t+5]\frac{L_0^F}{5\,\ell_0} \tag{1}$$

show how we can derive

$$N_0^F(t+5) = \sum_x N_x^F(t) b_x^F \tag{2}$$

where $b_x^F = \frac{1}{1+SRB}\frac{L_0^F}{2\ell_0}\big(F_x + s_x^F\, F_{x+5}\big)$.

Question 5

Take again the population that you used for Question 3. Project the population for 5 years ahead, but this time use matrix formulas. Compare your results with the projections that you obtained in Question 3. Are they the same?

Question 6

Project your chosen population by sex for $n = 20$ periods ahead and show your results.

Question 7

Following up on Question 6, include migration in your projection 20 period ahead. You can either use observed net migration counts (you can estimate them from the balancing equation of population growth) or the Rogers-Castro migration schedule (try to find reasonable data on total net migration counts).

Question 8

Project your chosen population by sex for $n = 20$ periods ahead, making time-specific assumptions about one or more demographic components.

Question 9

Derive the long-term growth rate and population distribution by eigendecomposing your Leslie matrix, and compare your results with those that you obtain from a long projection of your population. As a second step, modify your starting population by randomly re-assigning elements on $\boldsymbol{N}(t)$ to different age groups, and project the new population in the long term. Compute the long-term growth rate and distribution of this second population, and compare it to the first.

Bonus

For exercise 6 to 9, present your results with the aid of a shiny app or an animation