PEP-NET

European Commission

Predictive Epigenetics: Fusing Theory and Experiment

Title of the protocol: Human Population and Models

Authors & Affiliations: Karen A. Oliveira (GIE-CERBM) – ESR 8

Minimum Age: 14 years

Introduction:
Predicting human demographic behavior influences global decisions and policies. Thomas Malthus (from which the exponential model emerges) proposed that the population would grow in a geometric progression and the food production in an arithmetic progression. Thus, lack of resources, malnutrition, diseases would lead to a population decrease due to the death of part of the population [A. P. Dobson, 2008, Amano and Katayama, 2009].
Understanding how populations grow - and if they do - is also a way of predicting the future [Edelstein-Keshet, 2005, Murray, 2007].
Several factors over history influenced the size of the human population, from environmental tragedies like the eruption of Vesuvius to conquests like Genghis Khan in the 13th century. Or the pandemic we are currently living in.
Depending on the size of the impact, these events may influence the global demography. Considering mathematical modeling, we can determine growth models for the human population over the years.
In Fig. 1, we present the estimative made for the human population from 1000 to 1950. Note the visible increase in the size of this population over the years.

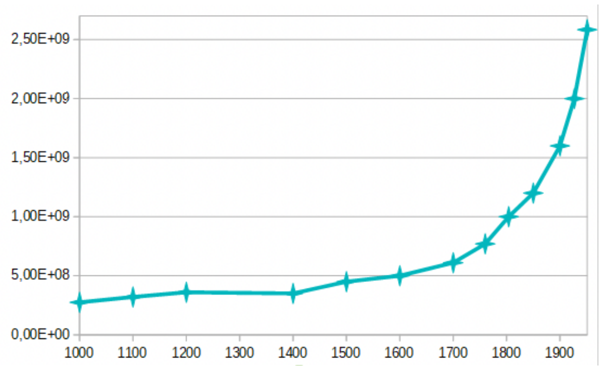


Fig. 1: Estimated human population from 1000 to 1950. Source: World Population by Year.

Is human population growth exponential?
What expression of growth can we have for this population?

Protocol:

To answer these questions, we consider the values from Fig. 1, Table 1 and we create a model that explains the population growth:

Table 1: Values of Fig. 1.

Year	Population
1000	275,000,000
1100	320,000,000
1200	360,000,000
1400	350,000,000
1500	450,000,000
1600	500,000,000
1700	610,000,000
1760	770,000,000
1804	1,000,000,000
1850	1,200,000,000
1900	1,600,000,000
1927	2,000,000,000
1951	2,584,034,261

We can solve this either mathematically or graphically:

1) **Exponential Function:**
We consider the formula for the exponential function, i.e.,

$$P(t) = ba^t;$$

Considering b = 1,600,000,000, we have:

$$P(27) = 2{,}000{,}000{,}000 = 1{,}600{,}000{,}000a^{27}$$

$$\frac{2{,}000{,}000{,}000}{1{,}600{,}000{,}000} = a^{27}$$

$$\frac{2.0}{1.6} = a^{27}$$

$$1.25 = a^{27}$$

$$\sqrt[27]{1.25} = \sqrt[27]{a^{27}}$$

$$a \approx 1.0083$$

Thus, our population model is $P(t) = 1{,}600{,}000{,}000 \cdot (1.0083)^t$. We validate our model:

Year	Population	Model
1900	1,600,000,000	1,600,000,000
1927	2,000,000,000	2,000,000,000
1951	2,584,034,261	2,438,923,112
1984	4,784,011,621	3,203,763,333

Does this model represent reality?
Why do we have such a high difference between the values measured and from our model?
Can you cite any historical event that could affect the population dynamics?
Do know how to modify this model in order to make it more realistic?

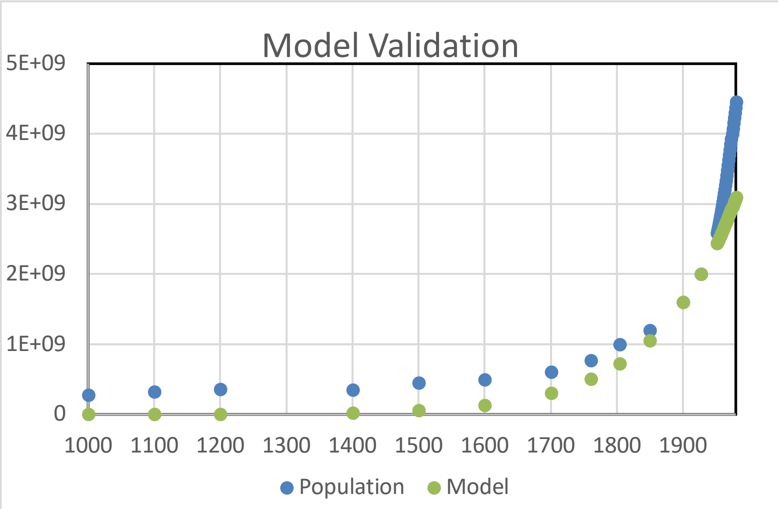


Fig. 2: Values from our model and the real data.

2) **Exponential Growth:** We consider the exponential function as a good approximation for the human population and let $t_0 = 1000$:

$$P(t) = P(t_0)e^{r(t-t_0)}$$

$$P(1927) = 2{,}000{,}000{,}000 = 275{,}000{,}000e^{r(1927-1000)}$$

$$\frac{2{,}000{,}000{,}000}{275{,}000{,}000} = e^{927r}$$

$$\frac{2{,}000}{275} = e^{927r}$$

$$\ln\left(\frac{80}{11}\right) = \ln(e^{927r})$$

$$1.9841 = 927r$$

$$r \approx 0.00214$$

Thus, our model in this case is:

$$P(t) = 275{,}000{,}000e^{0.00214(t-1000)}$$

We can also try to validate our model:

Year	Population	Model
1900	1,600,000,000	1,887,051,992
1927	2,000,000,000	1,999,297,400
1951	2,584,034,261	2,104,663,967
1984	4,784,011,621	2,258,669,264

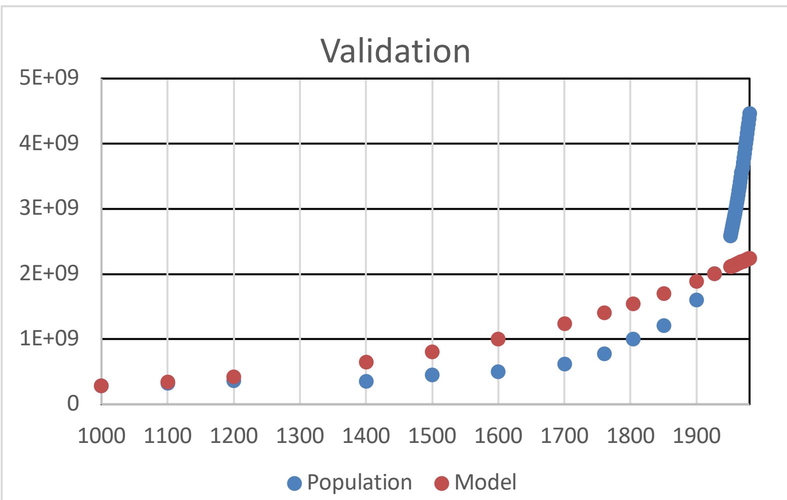


Fig. 3: Comparison between our model and real data.

Does this model represent reality?
Can you adapt this model to make it more realistic?
Which model seems more realistic?

We can also compare the two models we created:

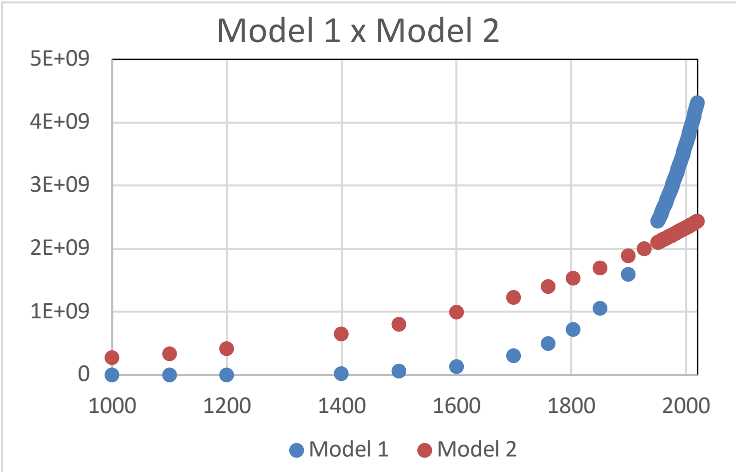
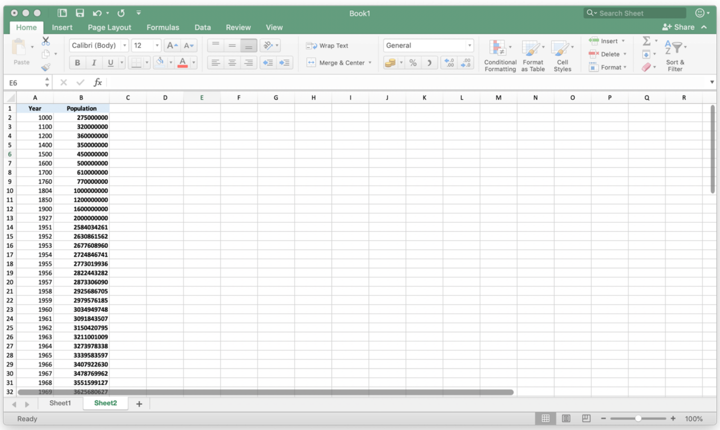


Fig. 4: Comparison between the two models

Can you explain the difference in obtained values from both models?

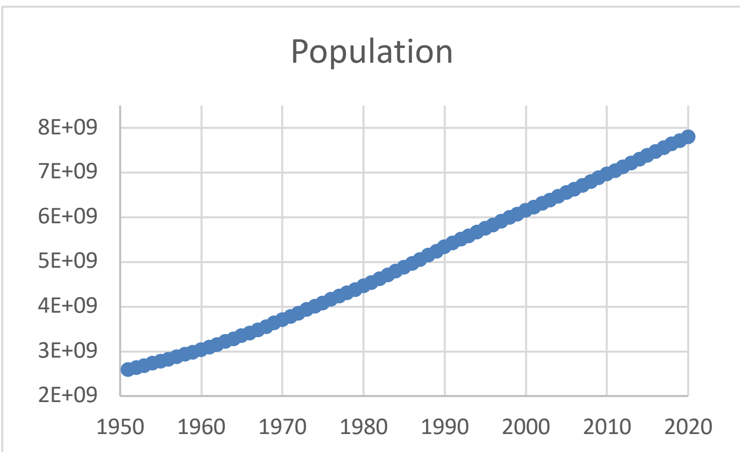
3) **Numerical Method:** This method needs at least a bit of knowledge of spreadsheets. Thus:

1. Pick the dataset from [WPb,2021] and put it on some spreadsheet;

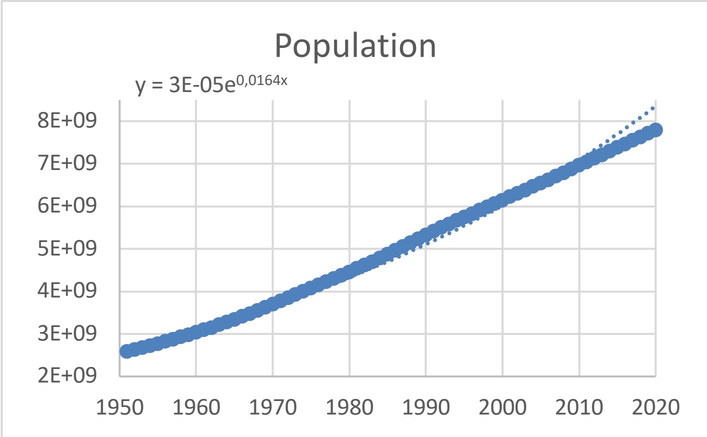


	A	B
1	Year	Population
2	1000	275000000
3	1100	320000000
4	1200	360000000
5	1400	350000000
6	1500	450000000
7	1600	500000000
8	1700	610000000
9	1760	770000000
10	1804	1000000000
11	1850	1200000000
12	1900	1600000000
13	1927	2000000000
14	1951	2584034261
15	1952	2630861562
16	1953	2677608960
17	1954	2724846741
18	1955	2773019936
19	1956	2822443282
20	1957	2873306090
21	1958	2925686705
22	1959	2979576185
23	1960	3034949748
24	1961	3091843507
25	1962	3150420795
26	1963	3211001009
27	1964	3273978338
28	1965	3339583597
29	1966	3407922630
30	1967	3478769962
31	1968	3551599127
32	1969	3625680627

2. Create a scatter plot from it;



3. Add a trendline and try to fit your data



Note that if we consider the human population after 1950, we have better results. This means that an exponential fitting is not realistic for this problem since the human population depends more than just itself to grow. Can you explain what changed over the last 70 years that you think affected this tendency?

Expected Difficulties and Questions:

1. Problems with the dimensions of the numbers worked (in billions of people). **How to help:** Explain scientific notation as base 10 exponents, explaining that it helps with the representation of either big and small numbers.
2. Difficulties to understand that ln e = 1. **How to help:** Explain to them more than once the correlation between the exponential and logarithmic function.
3. Use the raw data in a spreadsheet to build the model. **How to help:** Guide how to write functions in spreadsheets, insert graphics.
4. Understand how to represent the exponential growth function. **How to help:** Show that it is an exponential function based on the Euler number, e.
5. You will obtain different values if you try to write your model. This can raise doubts about which one is right. **How to help:** Explain that it depends on what data has been considered and what "close" errors are acceptable.

References:

- [WPb, 2021] (2021). World population by year.
- A. P. Dobson, B. T. G., editor (2008) Ecology of Infectious Diseases in Natural Populations. CAMBRIDGE UNIV PR.
- Amano, T. and Katayama, N. (2009) Hierarchical movement decisions in predators: effects of foraging experience at more than one spatial and temporal scale. Ecology, 90(12):3536-3545.
- Boyce, W. E. and DiPrima, R. C. (2005) Elementary differential equations and boundary value problems, eighth edition, volume 9. Wiley New York.
- Edelstein-Keshet, L. (2005) Mathematical Models in Biology. Cambridge University Pr.
- Murray, J. D. (2007) Mathematical Biology. Springer New York.