Phenomenological Models

> A struggle for existence inevitably follows from the high rate at which all organic beings tend to increase. Every being, which during its natural lifetime produces several eggs or seeds, must suffer destruction during some period of life, and during some season or occasional year, otherwise, on the principle of geometric increase, its numbers would quickly become so inordinately great that no country could support the product.
>
> *Charles Darwin, The Origin of Species, 1859*

Learning Objectives

After completion of this module, the student will be able to

1. define population growth rate, per capita growth rate, exponential growth and decay, logistic growth
2. use data to build a phenomenological model with a minimal number of parameters using the framework of differential equations
3. estimate model parameters

Knowledge and Skills

1. introduction to continuous time population models
2. average and instantaneous growth rates
3. doubling time of an exponentially growing population
4. population growth rate, per capita growth rate

Prerequisites

1. calculating percent changes
2. natural logarithm, exponential function
3. graphing in EXCEL
4. fitting a straight line to data points in EXCEL and displaying the equation
5. definition of the derivative using limits

Data on Population Growth

The U.S. Census Bureau keeps track of the world population, calculates annual growth rates, and uses models to predict future population sizes. For instance, the total midyear population for the world between 1950 and 2050 can be found at http://www.census.gov/ipc/www/idb/. This data is also

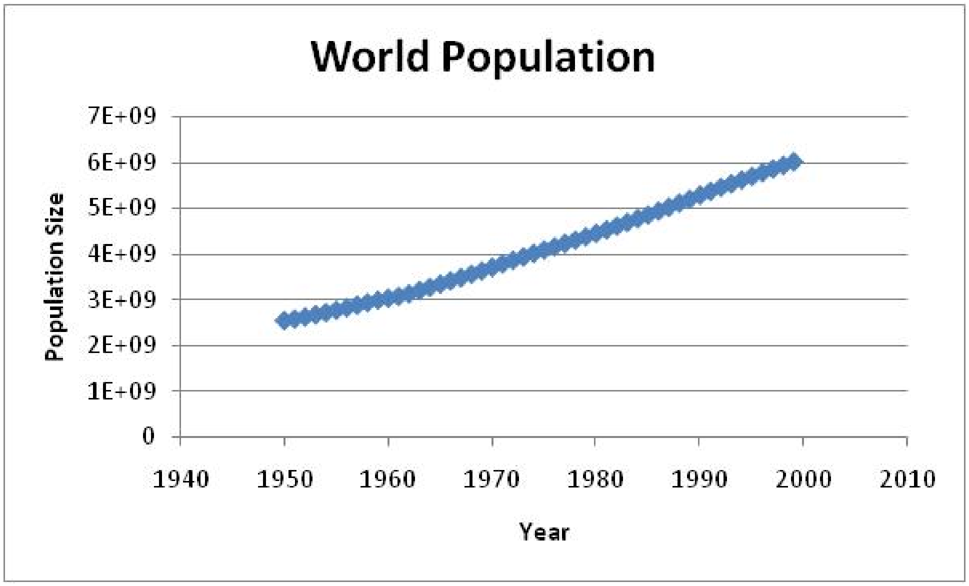

Figure 1: World population between 1950 and 1999. Data from U.S. Census Bureau.

available under the first tab in the accompanying spreadsheet (labeled World Population—the sheet is protected to avoid accidental overwriting of data). The population size between 1950 and 1999 is shown if Figure 1. We see a steady increase.

The annual growth rates in Column C of the spreadsheet are calculated using the formula

$$r(t) = \ln\frac{N(t+1)}{N(t)} \tag{1}$$

where t denotes the year, $r(t)$ the growth rate from midyear t to midyear t+1, $N(t)$ the population at midyear t, and ln the natural logarithm.

In-class Activity 1

1. Use Equation (1) to calculate the annual growth rate in percent from midyear t to midyear t+1 for the world population for $t = 1990, 1991, 1992, 1993$, and 1994:

Year	Midyear Population	Annual Growth Rate (%)
1990	5,282,371,928	
1991	5,365,708,797	
1992	5,448,725,522	
1993	5,529,987,425	
1994	5,610,065,463	
1995	5,690,982,026	

Compare your results with the annual growth rate that the U.S. Census Bureau calculated (see first tab in the accompanying spreadsheet).

2. The world population reached about 3 billion in 1959. How many years did it take the human world population to reach 6 billion, 9 billion? What is the percent increase during each of these time intervals?

3. Suppose a population doubles in size within a single year. What is the percent increase during that year? What is the average annual growth rate in % during that year based on Equation (1)?

Where Does the U.S. Census Bureau Formula Come From?

Some populations have non-overlapping generations, such as annual flowers or univoltine insects. Other populations have overlapping generations, such as humans or bacteria. In this module, we will only consider populations with overlapping generations.

In populations with overlapping generations, we model the population size in continuous time. The modeling framework is **differential equations**. If $N(t)$ denotes the population size at time *t*, then the change in population size during a time interval of length Δt, say between time *t* and time $t+\Delta t$ is given by $\Delta N = N(t+\Delta t) - N(t)$. The **average growth rate** during this time interval is then calculated as

(2) $$\text{average growth rate} = \frac{\Delta N}{\Delta t} = \frac{N(t+\Delta t) - N(t)}{\Delta t}$$

A population with overlapping generations is continually changing. To describe this change, we will calculate the **instantaneous rate of change**, that is, the rate of change during an **infinitesimal time interval**. To calculate this growth rate we take the limit as $\Delta t \to 0$:

(3) $$\text{instantaneous growth rate} = \frac{dN}{dt} = \lim_{\Delta t \to 0} \frac{N(t+\Delta t) - N(t)}{\Delta t}$$

You recognize this as the derivative of the population size $N(t)$ at time *t*. The derivative $\frac{dN}{dt}$ is the **population growth rate**. The **per capita growth rate** at time *t* is obtained by dividing the population growth rate by the population size

$$\text{per capita growth rate} = \frac{1}{N}\frac{dN}{dt}$$

Our first model says that the per capita growth rate is equal to a function, which we denote by $r(t)$. We thus have as our first model

(4) $$\frac{1}{N(t)}\frac{dN}{dt} = r(t)$$

Let's assume for the moment that the per capita growth rate is constant and let's denote it by *a*, that is, $r(t) = a$ for all *t*, then one can show using integral calculus that the solution of Equation (4) is

(5) $$N(t) = N(0)e^{at}$$

where $N(0)$ denotes the population size at time 0. We can confirm this by differentiating Equation (5). We find

$$\frac{dN}{dt} = a\underbrace{N(0)e^{at}}_{N(t)} = aN(t) \text{ and } N(0) = N(0)\underbrace{e^{(a)(0)}}_{=1}$$

In-class Activity 2

1. Assume that the human population has a constant per capita growth rate of 1.4%. Use Equation (5) to calculate how many years it will take the population to double in size?

2. To calculate the doubling time of a growing population with constant per capita growth rate, you divide the percent growth rate into 70. Apply this rule to check your answer in In-class Activity 2, Problem (1). Use Equation (5) to explain where this rule comes from.

We can now check why Equation (1) holds when the per capita growth rate is constant. Namely, looking at Equation (5), we see that

$$N(t+1) = N(0)e^{a(t+1)} \text{ and } N(t) = N(0)e^{at}$$

We find for the ratio $\frac{N(t+1)}{N(t)}$

$$\frac{N(t+1)}{N(t)} = \frac{N(0)e^{a(t+1)}}{N(0)e^{at}} = e^{a(t+1)-at} = e^{at+a-at} = e^{a}$$

Taking natural logarithms on both sides, we obtain

$$\ln\frac{N(t+1)}{N(t)} = a$$

This is Equation (1) in the case when the per capita growth rate is constant. To understand why Equation (1) can be used to estimate the per capita growth rate regardless of whether the per capita growth rate is constant, we need to be able to solve a differential equation by integration. [If you have not covered this in calculus, you can skip over the following calculation for now and go directly to Equation (6).] If the reproductive value or the per capita growth rate varies over time, the per capita growth rate can be approximated by solving the differential equation in Equation (4) from one time step to the next. Separating variables in Equation (4) yields

$$\frac{dN}{N} = r(t)dt$$

We will integrate this over the time interval [t,t+1]. Since the antiderivative of $1/x$ is $\ln x + C$, we find

$$\ln N(t+1) - \ln N(t) = \int_t^{t+1} r(s)ds$$

The left-hand side can be simplified using the rule $\ln x - \ln y = \ln\frac{x}{y}$. The right-hand side cannot be simplified without knowing more about the function $r(t)$. However, the right-hand side can be interpreted as the average value of the function $r(t)$ over the time interval $[t, t+1]$ using the formula of calculating the average value of a continuous function $f(x)$ over the interval $[a,b]$

$$f_{avg} = \frac{1}{b-a}\int_a^b f(x)dx$$

We therefore find that

(6) average per capita growth rate in $[t, t+1] = \ln\frac{N(t+1)}{N(t)}$

The formula in Equation (6) is the one used by the U.S. Census Bureau. It calculates the average per capita growth rate of the population for the time interval $[t, t+1]$. This is not the same as the instantaneous growth rate $r(t)$. In other words, when we use Equation (1), we find approximations for the instantaneous per capita growth rate $r(t)$ at time t by approximating $r(t)$ by the average per capita growth rate over the time interval $[t, t+1]$.

Back to Data

When looking at the world population between 1950 and 1999 (Figure 2), we see that the annual growth rate varies from year to year. We can make two observations for the time period between 1950 and 1999: (1) The world population increased every year during this time interval; and (2) there are periods during this time interval when the rate of growth is increasing and when it is decreasing.

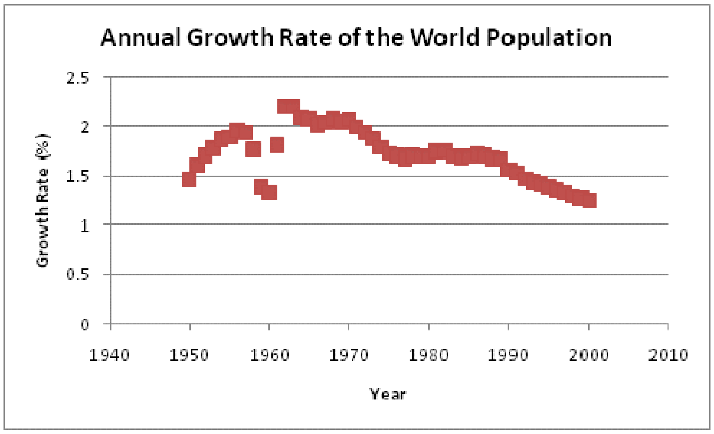

Figure 2: Annual growth rate between 1950 and 1999 of the World Population. (Source: U.S. Bureau of Census)

In-class Activity 3

1. As long as the growth rate is positive, the population size increases, regardless of whether the growth rate is increasing or decreasing. What is the difference in the graph of a population size versus time when the population is increasing at an increasing rate and increasing at a decreasing rate?

2. How can you extend this to populations that are decreasing in size?

3. Can you come up with other examples where a quantity is increasing/decreasing at an increasing/decreasing rate?

	Increasing rate	Decreasing rate
Increasing quantity		
Decreasing quantity		

The dynamics of the human population are quite complex. Economic and political factors play a significant role and are difficult to model. For instance, the dip in the annual growth rate from 1959-1960 was due to the Great Leap Forward in China where a combination of natural disasters and decreased agricultural output during social upheavals caused a sharp rise in death rates and a decline in its birth rate.

To avoid the complexities of human population growth, we will now turn to a much simpler organism that can grow quickly in a controlled environment, namely yeast that grows in a medium at a constant temperature in the lab. We will start with some background reading on modeling in population biology.

Background Reading on Modeling in Population Biology

W.M.Getz. 1998. An introspection on the art of modeling in population ecology. *BioScience* 48(7), 540-552. (Read pages 540-547: introduction, historicity in modeling, kinds of models, resolution and focus, a much-ignored phenomenological principle)

In-class Activity 4

Discuss the following quotes from the paper by Getz:

P 545: "Another way to classify models is to categorize them as phenomenological, empirical, or mechanistic. A phenomenological approach focuses on a direct characterization of the holistic properties of the population without elaborating the underlying mechanisms that might be responsible for these properties."

P 545: "Mathematical models, like other types of models, are always much less than, and of a substance much different from, the things they represent. A real art in ecological modeling is to select the appropriate level of resolution to convey a particular property of a real system."

P 544: "[…] historicity in modeling represents conservative, and even repressive, intellectual forces that are antithetic to the art of modeling in population ecology."

The Yeast Experiment

A fresh medium was prepared and inoculated with a small number of yeast cells that were grown overnight in the same type of medium. A spectrophotometer was used to measure the optical density of the medium every hour, which serves as a proxy for cell density. The second sheet of the accompanying spreadsheet contains the hourly measurements of OD600 from four replicates of this experiment. The date was obtained by Dr. Brett Couch, a teaching assistant professor in the Biology Program at the University of Minnesota Twin Cities.

Task 1: Plot the optical density as a function of time for all four replicates in a single graph. The data of this experiment can be found in the second sheet of the accompanying spreadsheet (labeled Raw Data). The cells in that sheet are protected to avoid accidental overwriting of values.

1. Copy the data from the second sheet (labeled *Yeast Raw Data*) into the third sheet (labeled *Task 1*). The third sheet is unprotected so that you can work with the data, for instance to create a plot.
2. To create the plot in this task, highlight the cells that contain the time points (first column) and the data (next four columns) in the second sheet. Click on the **Insert** tab. Choose **Scatter** in the **Charts** group, and click on the **Scatter with Smooth Lines and Markers** option. This will plot optical density as a function of time for all four replicates in a single graph.
3. To label the data, click on the graph. This opens the **Chart Tools**. Click on the **Design** tab and on **Select Data** in the **Data** group. The **Select Data Source** window opens that allows you to make changes to the selected data. We will use it to label the data. Click on **Series 1** in the window. Click on **Edit**. Type *Flask #1* in the **Series name** field, click OK in the **Edit Series** window. Use this procedure to label Series 2-4, which correspond to the data from Flasks #2-#4. When you are done, click **OK** in the **Select Data Source** window,
4. To label the graph, click on the graph, and then on the **Layout** tab in the **Chart Tools**. In the **Labels** group, click on **Chart Title** and choose the **Above Chart** option. Type *OD 600 Measurements*. Hit **Enter** on your keyboard. Next, click on **Axis Titles** in the **Labels** group. Click on the **Primary Horizontal Axis Title** and choose the **Title Below Axis** option. Type *Time Points [hrs]* and hit **Enter** on your keyboard. To label the vertical axis, click on **Axis Titles** in the **Labels** group. Click on the **Primary Vertical Axis Title** and choose the **Rotated Title** option. Type *Population Size* and hit **Enter** on your keyboard. Your graph is now ready.
5. Copy the labeled graph into a Word file. Write a paragraph describing the growth pattern.

Estimating the Rate of Growth of the Yeast Population

Let's find the per capita growth rate using Equation (6). The yeast population in our experiment has overlapping generations and so we can use a similar approach as for the world population data.

Task 2: Go to the fourth sheet in your spreadsheet (labeled *Task 2 and Task 4*). The raw data of the first flask are copied on this sheet. Calculate the per capita growth rate $r(t) \approx \ln \frac{N(t+1)}{N(t)}$ [see Equations (6)], and produce a graph (using scatter plots "Scatter with Only Markers") to graph the per capita growth rate $r(t)$ as a function of *t*. Describe what you see.

When yeast is introduced into a new medium, there is often a delay before reproduction begins. In addition, the accuracy of data is not very good when the density is low. To build a model of growth, we will therefore eliminate this initial period. Looking at your data, which points should we ignore?

It is reasonable to try out whether the per capita growth rate can be described as a function of population size. In other words, we will try the model

(7) $$\frac{1}{N(t)} \frac{dN(t)}{dt} = r[N(t)]$$

Task 2 (ctd.): To build the model, graph (using scatter plots "**Scatter with Only Markers**")the per capita growth rate at time *t* as a function of population size at time *t*, namely $N(t)$ starting at time $t = 10$. (This eliminates the initial period where cells are at low density.) Describe what you see.

Determine the equation of the function $r(N)$ as a function of N using the Trendline Tool:

1. Click on the data points in your graph. The **Chart Tools** group appears. Click on the **Layout** tab in the Chart Tools group. Choose **Trendline** in the **Analysis** group. Click on "**More Trendline Options...**" A window appears. Since we wish to fit a straight line, choose **Linear** for the **Trend/Regression Type**. Check the boxes for **Display Equation on chart** and. Click on **Close**.
2. The equation for the fitted straight line is displayed on the graph. In your lab report, include the scatter plot together with a figure legend and the equation for $r(N)$. Note that the equation in your figure is given in terms of the variable *x*. In your report, use the variable *N*.

Logistic Growth

In Task 2, you derived a phenomenological model for the growth of the yeast cultures. The model you derived is of the form

(8) $$\frac{dN}{Ndt} = b - aN$$

where the parameters a and b are positive and were estimated using the Trendline Tool. [In Equation (8), we omitted the dependence of the population size on t. This is customary in the biology literature.]

The next step is to interpret the parameters. We have two parameters. To interpret them, we look at per capita growth at very low densities and at long-term behavior. To find the growth rate at very low densities, we set $N = 0$ on the right-hand side of Equation (8), which then simply becomes b. That is, the parameter b describes the per capita growth rate at very low densities. To describe the long-term behavior, we set the right-hand side of Equation (8) equal to 0, indicating that after a long time, there might be no change in the population size. We find that in this case, $N = b/a$.

Looking back at the yeast data, we saw that the population density seemed to approach a constant value. We call this value the carrying capacity and can now identify it with the quantity b/a, which is often denoted by the capital letter K. The per capita growth rate at very low densities is called the intrinsic rate of growth and is typically denoted by the lower case letter r. Using the new parameters r and K, we can rewrite Equation (8)

(9) $$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

This model is called Logistic Growth Model.

Task 3: (a) Derive Equation (9). (b) What are the estimated values for the parameters r and K?

A Different Approach (not needed for Project 1 below)

The approach we chose to model population growth was motivated by the U.S. Bureau of Census' formula for annual growth rate. We were able to find a simple model since the per capita growth rate was a linear function of population size. A different approach uses the definition of the derivative.

Equation (9) is of the form

$$\frac{dN}{dt} = f(N)$$

where $f(N)$ denotes the population growth rate. Using the definition of the derivative, we can write this as

$$\lim_{h \to 0} \frac{N(t+h) - N(t)}{h} = f(N)$$

Thus, we can approximate $f(N)$ as

$$f(N) \approx \frac{N(t+h)-N(t)}{h}$$

Plotting $\frac{N(t+h)-N(t)}{h}$ as a function of *N* might reveal a pattern that we recognize as one of the standard functions, for instance, a linear function or a quadratic function, we can use the Trendline Option in EXCEL to fit a function to the data and estimate the parameters. An important task then remains to give biological meaning to the parameters.

We can use this method also to estimate the relative growth rate $g(N) = \frac{1}{N}\frac{dN}{dt}$. Namely,

$$g(N) \approx \frac{1}{N(t)}\frac{N(t+h)-N(t)}{h}$$

Plotting the right-hand side as a function of *N* might reveal a simple function that we can fit to the data and estimate the parameters.

Task 4: Since we already know that the relative growth rate of the yeast in our experiment fits a simple model, namely

$$\frac{1}{N}\frac{dN}{dt} = r\left(1-\frac{N}{K}\right)$$

which is a straight line, we will use the data to estimate the parameters *r* and *K* of the model in this way. The data in the Flask #1 experiment were collected every hour, so $h=1$. In the spreadsheet under the Task 2 and Task 4 tab, calculate

$$\frac{N(t+1)-N(t)}{N(t)}$$

and plot this as a function of $N(t)$ (use an appropriate graph). Fit a straight line. Determine the parameters *r* and *K* and compare your estimates to the estimates you found in Task 3.

Summary

Suppose we observe a variable *V* changing as a function of *x*. We collect data for $x = x_0, x_0+1, x_0+2, \ldots$ and wish to build a model that would describe how *V* changes with *x*. We learned two approaches: the first approach models the relative rate of change

$$\frac{1}{V}\frac{dV}{dx} = g(V)$$

where we graph $\ln\frac{V(x+1)}{V(x)}$ as a function of $V(x)$. The second approach relies on the definition of the derivative and is suitable for modeling both the rate of change and the relative rate of change of a quantity

$$\lim_{h\to 0}\frac{V(x+h)-V(x)}{h} = f(V)$$

or

$$\frac{1}{V(x)}\lim_{h\to 0}\frac{V(x+h)-V(x)}{h} = g(V)$$

This yields the approximations

$$f(V) \approx \frac{V(x+h)-V(x)}{h} \text{ or } g(V) \approx \frac{1}{V(x)}\frac{V(x+h)-V(x)}{h}$$

For the rate of change $f(V)$ or the relative rate of change $g(V)$.

Project 1

This project continues the discussion on the world population data. The first tab in the accompanying spreadsheet lists the world population for the years 1950 to 1999. Figure 2 in this module shows how the annual per capita growth rate changed over time during this period. Use the data to build a model that would allow you to predict the world population for the years 2000 to 2050. What are your assumptions? Compare your predictions to those of the U.S. Census Bureau (http://www.census.gov/ipc/www/idb/worldpop.html). Present your findings as a poster.

Source: U.S. Census Bureau (http://www.census.gov/ipc/www/idb/worldpop.html)

Project 2

In your EXCEL file under the tab *Carlander* you will find data for a population of shallow-water ciscoes (*Coregonus artedi*). The data gives the average length for this fish as a function of age. Use the data to build a phenomenological model. Present your findings as a poster.

Source: Carlander, K.D. 1950. *Handbook of freshwater fishery biology*. Wm. C. Brown Co., Dubuque, Iowa. 281 pp.

Project 3

In your EXCEL file under the tab VMRC bluefish you will find data for bluefish. The data is explained in the file bluefish1999.pdf. Build a phenomenological model that relates the average length for this fish to age. Present your findings as a poster.

Source: 1999 VMRC Annual Report
http://www.odu.edu/sci/cqfe/Research/Chesapeake%20Bay/Bluefish/Bluefish.htm

Additional Resources

Design of Scientific Posters:
http://www.hhmi.org/coolscience/resources/SPT--FullRecord.php?ResourceId=29
Statistics for Journalists:
http://www.robertniles.com/stats/