Dylan Zwick
Math 2280
Spring, 2008
University of Utah

Population Prediction Project Demonstration Writeup

In this project we will investigate how well a logistic population model predicts actual human populations, using both United States census data, and world census data.

Now, a logistic population model is the solution to an ordinary differential equation of the form:

$$\frac{dP}{dt} = aP + bP^2 \qquad P(0) = P_0$$

So, there are three "unknown" constants in this model: the constants a and b in our differential equation, and our initial population. If we rewrite our equation in the form:

$$\frac{1}{P}\frac{dP}{dt} = a + bP$$

we see that if we graph the values of our population vs. the value of $\frac{1}{P}\frac{dP}{dt}$ then if our model is correct we should get a linear relation with slope b and y-intercept a. Of course, our model won't be completely accurate, but we can use a linear regression line to find a "best fit" regression curve for our linear relation, and use the slope and y-intercept values of this regression curve to estimate the parameters in our differential equation, and examine how well these estimates, combined with our logistic population model, predict the populations.

The next question is, how do we calculate, or estimate, the value of $\frac{dP}{dt}$. A reasonable way to estimate this value is by taking, at every time, the population of the measurement immediately before it, and the population of the measurement immediately after it, and draw a straight line between these two values, and use the slope of this line are our estimate for $\frac{dP}{dt}$. As a formula:

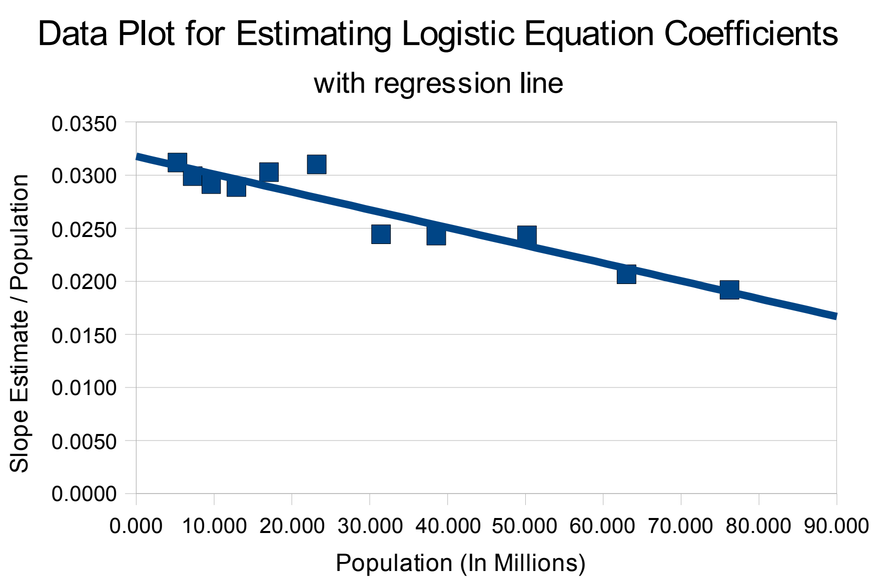
$$\left(\frac{dP}{dt}\right)_i \approx \frac{(P_{(i+1)} - P_{(i-1)})}{(t_{(i+1)} - t_{(i-1)})}$$

Investigation A

Using the method described above along with the times and population numbers from our textbook we can construct the following table:

Year	Population (In Millions)	Slope Estimate	Slope Estimate / Population
1790	3.929		
1800	5.308	0.166	0.0312
1810	7.240	0.217	0.0299
1820	9.638	0.281	0.0292
1830	12.861	0.371	0.0289
1840	17.064	0.517	0.0303
1850	23.192	0.719	0.0310
1860	31.443	0.768	0.0244
1870	38.558	0.937	0.0243
1880	50.189	1.221	0.0243
1890	62.980	1.301	0.0207
1900	76.212	1.462	0.0192
1910	92.228		

Constructing a plot of the second and fourth columns of this data, with population on the horizontal, and slope estimate / population on the vertical we get the following data plot:



The equation for the regression line above is:

$$-0.0001676x+0.03176$$

with an R^2 value of 0.889.

Now, if we solve the logistic growth equation given at the beginning of this writeup we get:

$$P(t)=\frac{aP_0}{((a+bP_0)e^{(-at)}-bP_0)}$$

which, when we plug in our values for a and b, along with our initial population of 5.308 million gives us:

$$P(t)=\frac{.16858}{(.03087e^{(-.03176t)}+.00089)}\quad .$$

If we use this model to predict the U.S. Population from 1900 to 2000 we get, compared to the actual populations:

Year	Census Population (In Millions)	Estimated Population (In Millions)
1900	76.212	77.370
1910	92.228	92.212
1920	106.022	107.178
1930	123.203	121.536
1940	132.165	134.668
1950	151.326	146.164
1960	179.323	155.847
1970	203.302	163.743
1980	226.542	170.013
1990	248.710	174.888
2000	281.422	178.616

These are pretty decent predictions, except when we get to around 1950, then they start to diverge. The reason for this has much to do with modern medicine, and a little less to do with the agricultural revolution, but basically modern science changed the carrying capacity of the United States in the second half of the 20th century, and so the old population model no longer applied.

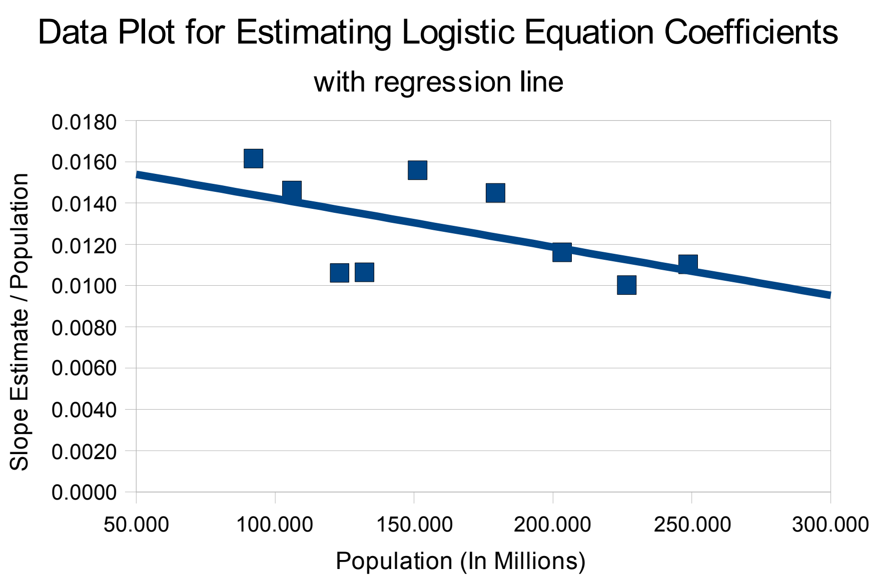
Investigation B

Here we repeat the methods we used in Investigation A, only we just use 20th-century population data and use 1900 as our initial population.

So, using just 20th-century population data we construct the table:

Year	Population (In Millions)	Slope Estimate	Slope Estimate / Population
1900	76.212		
1910	92.228	1.491	0.0162
1920	106.022	1.549	0.0146
1930	123.203	1.307	0.0106
1940	132.165	1.406	0.0106
1950	151.326	2.358	0.0156
1960	179.323	2.599	0.0145
1970	203.302	2.361	0.0116
1980	226.542	2.270	0.0100
1990	248.710	2.744	0.0110
2000	281.422		

and we get the regression plot:



with regression line equation:

$$-.0000235x+.01657$$

with an R^2-value of .2845. We note this value is much less than our earlier value, because of the change in the appropriate population model around the middle of the 20th-century.

Now, using these parameter and an initial population of 76.212 million in 1900 we get:

$$P(t)=\frac{53737}{(629e^{(-.01657t)}+76)}$$

Now, using this logistic equation we get the population table:

Year	Population (In Millions)	Logistic Population Estimate
1900	76.212	76.22
1910	92.228	88.25
1920	106.022	101.86
1930	123.203	117.17
1940	132.165	134.28
1950	151.326	153.23
1960	179.323	174.05
1970	203.302	196.69
1980	226.542	221.06
1990	248.710	246.98
2000	281.422	274.23

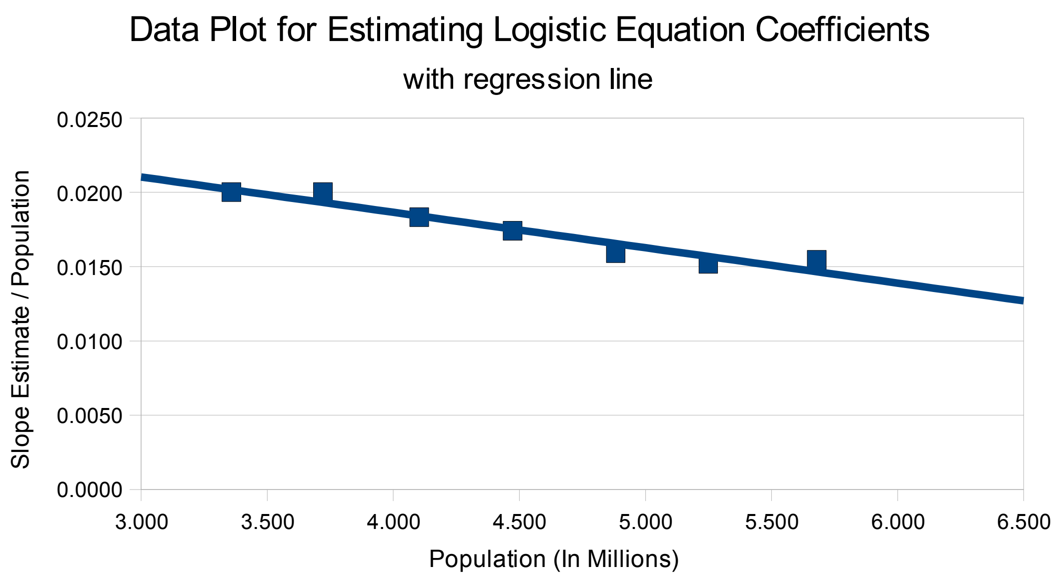
which we can see fits the actual population data much better, although it frequently underpredicts the actual population.

Investigation C

For this investigation we follow exactly the same procedure that we used for investigations A and B, although instead of using U.S. population data, we use U.N. world population data from the second half of the 20th century. Using this data we construct our table:

Year	Population (In Billions)	Slope Estimate	Slope Estimate / Population
1960	3.049		
1965	3.358	0.067	0.0200
1970	3.721	0.075	0.0200
1975	4.103	0.075	0.0183
1980	4.473	0.078	0.0174
1985	4.882	0.078	0.0159
1990	5.249	0.080	0.0152
1995	5.679	0.088	0.0155
2000	6.127		

with regression line plot:



and regression equation:

$$-.002379\,x+.02817$$

with an R^2-value of .9265.

Using these parameters in our logistic model, along with an initial population of 3.049 billion in 1960 we get the equation:

$$P(t)=\frac{36.1}{(8.79\,e^{(-.0281\,t)}+3.05)}$$

Using this logistic model we would predict a world population in 2025 of 8.085 billion, which is very close to the U.N. estimate of 8.177 billion. So, using some simple differential equations we can get the same predictions as the United Nations.