Contents

Fourth International Derive TI-89/92 Conference

Liverpool John Moores University, July 12 – 15, 2000

Using All of the Tools of Modelling: Modelling Populations

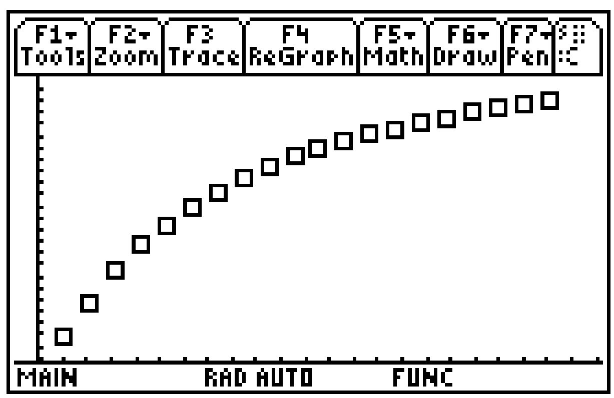
Wade Ellis, Jr.
West Valley College, USA
wade_ellis@wvmccd.cc.ca.us

Carl Leinbach
Liverpool John Moores University
&
Gettysburg College, USA
leinbach@cs.gettysburg.edu

Collecting Data and Curve Fitting

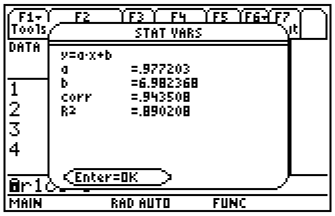
The TI-89 in conjunction with the CBL2 allows you to collect data and then to analyse it with standard statistical curve fitting tools. As an introduction to curve fitting tools we look at data collected from a temperature probe. An example is shown below. For the first screens the column on the right is time in seconds, the second is the temperature of a cooling body, and the third column is this temperature less the ambient temperature. A data table and its graph of information collected from a temperature probe that had been immersed in ice water and the grasped by the finger tips for 19 seconds recorded from 1 to 20 are displayed below.

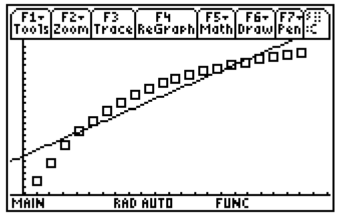
Time	Temp.	Time	Temp.
1	3.01	11	19.61
2	6.01	12	20.29
3	8.66	13	20.86
4	10.95	14	21.42
5	12.80	15	21.87
6	14.49	16	22.33
7	15.79	17	22.77
8	16.96	18	23.11
9	17.89	19	23.45
10	18.81	20	23.78



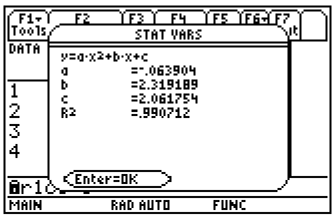
We display below the polynomial and regression coefficients and the function graphs for linear, quadratic, cubic, and quartic the regression as computed on the TI-89.

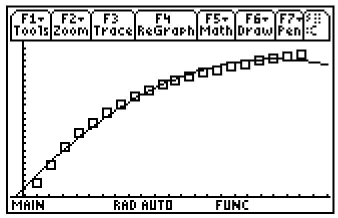
Linear Function



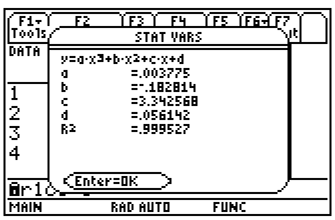


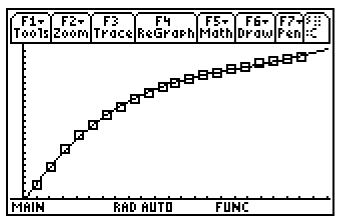
Quadratic Function



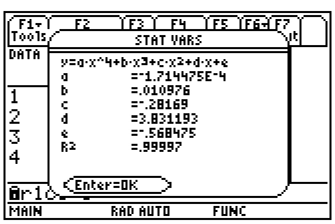


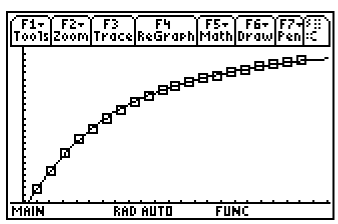
Cubic Function





Quartic Function





Although these tools can develop good models, they do not give the modeler much insight into why the resulting models work.

Building a Model Using Difference Equations

Difference equations can be used to develop models based on natural principles that give more insight into the phenomena that we are investigating. An underlying idea that can be used in develolping such models is:

Fourth International Derive TI-89/92 Conference

Future = Present + Change

This equation can be rendered as:

$$u(n) = u(n-1) + change$$

using the TI-89's SEQUENCE feature where

$$change = k \cdot (AmbTemp - u(n-1))$$

We develop this model for the temperature probe data based on the idea that the change in temperature is proportional to the difference between the current temperature of the probe and the ambient temperature of the room.

For the temperature probe data where the ambient temperature was known to be 28.6 degrees centigrade, we have the following equations:

$$u(n) = u(n-1) + k \cdot (28.6 - u(n-1))$$

There are several ways of estimating k. Algebraically, we have that

$$k = \frac{u(n) - u(n-1)}{28.6 - u(n-1)} \qquad .$$

Values for k can be computed in the data table using a user-defined forward difference function (fd(*list*=shift(*list*,1)-*list*) as shown in the left display below.

DATA	c1	c2	c3
1	1	3.01	.11723
2	2	6.01	.11731
3	3	8.66	.11484
4	4	10.95	.10482

c3=fd(c2)/(28.6-c2)

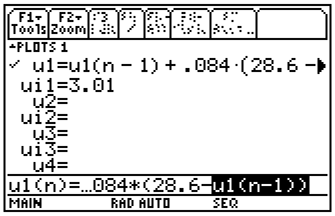
DATA	c2	c3	c4
1	3.01	.11723	.08391
2	6.01	.11731	
3	8.66	.11484	
4	10.95	.10482	

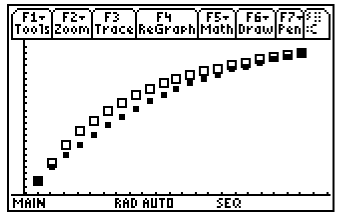
c4=mean(c3)

Note: The last value in c3 is undef. This value must be removed by unlocking c3 and then deleting the undef entry.

We may chose to average the values of in the c3 column as shown in the right hand display above. We display below the difference equation with initial value on the left and both the graph of the temperature probe data and the graph of the difference equation (black squares) on the right.

Fourth International Derive TI-89/92 Conference





As we mentioned before, there are several other methods for determining *k*. Let’s explore some of them together.

Workshop Exercises

1. Are the 19 data table values of *k* about the same?

2. What factors do we need to take into account in estimating *k*?

3. What other methods for estimating *k* are possible?

4. Which method is best?

We can compare the difference equation model to the polynomial models developed using curve fitting methods.

Workshop Exercises

1. How do the regression models compare with each other and with our difference equation model?

2. Which method is best and why?

From Discrete to Continuous
(from Joseph Fiedler of California State University)

We now look at moving from our recursion formula (difference equation) model to a differential equation model. In the temperature probe data, the recursion for one second is

$$u(n) = u(n-1) + 0.084 \cdot (28.6 - u(n-1))$$

Define $f(u) = u + 0.084 \cdot (28.6 - u(n-1))$. If we halve the time interval, then we would like to have two applications of the function (1/2+1/2 = 1 in the recursion process) to have the same result as one application of the original recursion formula which is based on a time interval of 1 second.

Solving the $f(f(u)) = u + 0.084 \cdot (28.6 - u(n-1))$ where we have omitted the subscripts gives two values of k. We can look at this in a factored form

factor(f(f(u))-u + 0.084*(28.6-u),u)=0

We can look at the same process for dividing the time interval by 2, 3, and 4.

factor(f(f(u))-u + 0.084*(28.6-u),u)=0

factor(f(f(f(u)))-u + 0.084*(28.6-u),u)=0

factor(f(f(f(f(u))))-u + 0.084*(28.6-u)=0

etc.

We begin to see a pattern.

(28.6-u)*((1-k)^2-1+0.084 = 0

(28.6-u)*((1-k)^3-1+0.084 = 0

(28.6-u)*((1-k)^4-1+0.084 = 0

etc.

The general formula for the roots might be proved by induction using the TI-89 if you should desire. Consequently, the constant for the time $1/n$, i.e., n iterations per second is a solution to the equation

$$(1-k_n)^n = 1 - 0.084$$

or

$$k_n = \left(1 - 0.916^{1/n}\right)$$

It is interesting to note that when we compute k_n for n even, we get two roots. For the sequence u_n, the larger root in each such pair of roots gives alternating convergence to 28.6.

To begin the transition from the discrete to the continuous, we now reinterpret the recursion formula as a difference equation. We recall that

$$u_n = u_{n-1} + 0.084 \cdot \left(28.6 - u_{n-1}\right)$$

or

$$u_n - u_{n-1} = 0.084 \cdot \left(28.6 - u_{n-1}\right)$$

and we can write

$$\Delta u = k_n \cdot (28.6 - u)$$

Dividing both sides of this equation by the change in time, , gives

$$\frac{\Delta u}{\Delta t} = \frac{\overbrace{k_n}^{\kappa_n}}{1/n} \cdot (28.6 - u)$$

For time step $1/n$, the appropriate coefficient of $(28.6 - u)$ is given as

$$\kappa_n = \frac{1 - 0.916^{1/n}}{1/n}$$

So the rate of change is

$$\frac{\Delta u}{\Delta t} = \frac{1 - 0.916^{1/n}}{1/n} \cdot \left(28.6 - u\right)$$

As $n \to \infty$, we have that $\kappa_n \to \kappa$, the coefficient of $(28.6 - u)$ in the associated differential equation. Thus, we take the limit of

$$\frac{1 - 0.916^{1/n}}{1/n}$$

as n approaches ∞ with the TI-89 to compute κ. The value of

limit((1-0.916^(1/n)),n, ∞)

is approximately 0.088. Of course, this computation uses l'Hopsital's rule. And so we have,

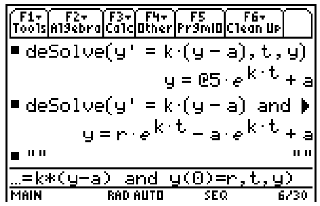
$$\frac{du}{dt} = \kappa \cdot \left(28.6 - u\right)$$
$$\frac{du}{dt} = 0.088 \cdot \left(28.6 - u\right)$$

is the continuous equivalent to the recursion formula. Thus, we have completed the transition from the discrete model to a continuous one.

You may wish to look at how this model fits the data.

The Continuous Case – Differential Equations

We can make the transition from the difference equation to the differential equation (Newton’s Law). The following screen shows the Differential Equation solver that is available on the TI-89 and TI-92 plus calculators



The screen shows a general solution for Newton’s Law of Cooling. The first equation assumes no initial condition. For the second application we assumed that $y(0) = r$.

To develop the equation for our particular situation, we will look at a sequence of difference equations and the sequence of associated k values for smaller and smaller time intervals to arrive at an estimate for the constant, k, in the differential equation.

We work on this development together in the workshop.

Modelling Populations (Using What We Learned From the Heat Equation)

We look at the more complicated situations involving population. We will start with a data set that you will be given and develop the models starting with the difference equation model and working to a differential equations model.

Moving on to Logistic Growth and Harvesting

Let’s look at a population that is living in a situation of limited resources. This is model that was developed by Volterra and refined by the mathematical biologist, Lotka in the 1940’s and 1950’s. The basic assumption of this model is that a population grows at a rate proportional to its’ present size (the standard assumption), but this growth also has an inhibiting factor that is proportional to the number of pairings within the population that are competing for the resource.

$$\textit{population growth} = k*\textit{population size} - c*\textit{competition}$$

If we let $y(t)$ denote the population size at time, t, then the first term on the right is given by

$$k*y(t)$$

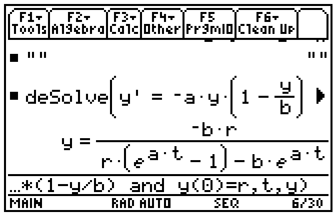
As we mentioned, the second term is determined by the number of pairings within the population, or

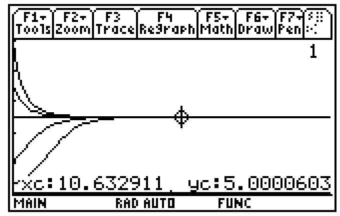
$$c*\frac{y(t)*(y(t)-1)}{2}$$

Gathering like terms, we arrive at the following, generally well known, equation

$$y' = a*y(t)*(1-\frac{y(t)}{b})$$

Using the **desolve(** operator on the TI-89 or TI-92 plus and assuming that $y(0) = r$, we see





On the left we show this solution with $a = 1$, $b = 5$, $r = \{0.5, 2, 5, 7, 10\}$, and $0 \leq t \leq 20$. Note that all of the solutions rather quickly stabilize at $y = 5$.

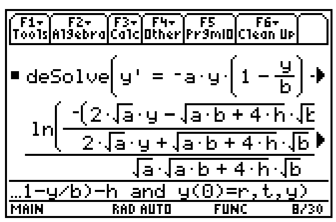
We now introduce a harvesting factor. This changes our model to:

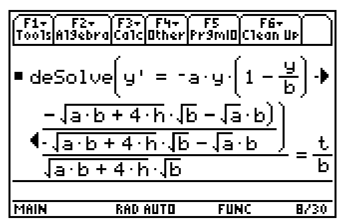
*population growth = k*population size – c*competition – harvesting*

If we assume the harvesting is at a constant rate, h, the differential equation becomes

$$y' = a * y(t) * (1 - \frac{y(t)}{b}) - h$$

Note that two screens are required to show the analytical solution and we still do not have the expression for y.



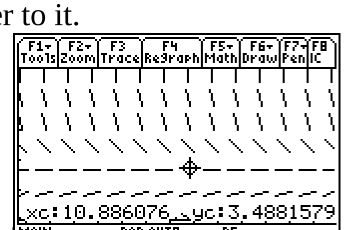


Providing specific values for the parameters does not make the situation much better. For example, try $a = 1$, $b = 5$, $h = 1$, and $r = 1.2$. There must be a better way.

Numerical Solutions to Differential Equations

Now we bring the discussion almost full circle. We begin with a differential equation model and use a numerical approximation technique. For this workshop instead of using the Euler, or linear approximation, we will use a version of the Runge-Kutta approximation techniques. Under the MODE screen 1, item 1, Graph, we choose option 6, DIFF EQUATIONS. In the Y = editor choose F1 option 9, Format. Be sure that Solution Method is set to RK and that Fields is set to Slope.

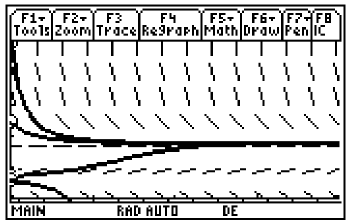
There are several nice things about this mode of display. First note that we can see a "flow" to the model as a whole. We are not stuck with looking at one equation. This display has a dynamic character to it.



Contents

At first appearance, the graph seems to be the same as one might expect for the standard logistic model except that the equilibrium value is a bit lower. It was 5 for the standard logistic model and appears to be about 3.5 here.

In addition to seeing the flow, we can look at what happens with the model given different assumptions about the initial conditions. In the following figure, we see curves for $y(0) = 1.2$, 1.5, 7, an 10



Now we spot a major difference. All of the solutions behave as expected except the one for $y(0) = 1.2$. For that one the population falls off to 0, i.e., becomes extinct. We have made a valuable discovery! However, this leads to a more important question: how is the level that leads to extinction related to the harvesting rate?

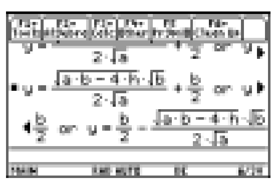
We look at this question now. The system is in equilibrium when $y'(t) = 0$. For our equation, this means

$$a * y * (1 - \frac{y}{b}) - h = 0$$

or

$$y^2 - b\ y^2 - b * y + \frac{b * h}{a} = 0$$

Solving this symbolically on the TI-89 we have



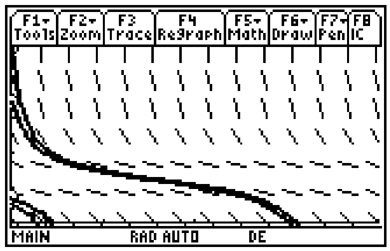
Note that when $h = 0$, we have that $y = b$ or $y = 0$. This is the standard logistic model. For the case when $a = 1$, $b = 5$, and $h = 1$ we have that the equilibrium values are at

$y = 3.618034$ and $y = 1.3820$

This is somewhat consistent with our previous investigation.

There is another observation that we can make from the slope field picture of our equation. The larger of the equilibrium values is a "stable" equilibrium. This means that any initial value for y within a given range of this equilibrium value will tend towards this value. The other equilibrium value is an "unstable" equilibrium. In this case, any initial value for y other than one which is absolutely on this equilibrium value will yield a curve which tends away from the equilibrium either to the other equilibrium value or towards extinction.

There is yet one other observation. When $h > \frac{ab}{4}$, then there are no real roots to the equation. Thus, there is no equilibrium. What happens in this case. When a = 1 and b = 5, then $\frac{ab}{4} = 1.25$. Let's see what happens for h = 1.4 and the same initial values for y.

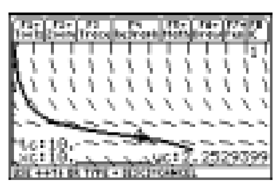


It appears that no matter where we start, the population is doomed to extinction! This is the sad message of over harvesting a population that has limited resources to support its existence.

Using the Model for Formulating Strategies

On of the strengths of having a modelling tool is the ability to ask 'What If ?' type of questions. For example, what if in the tenth year of the decline of the population in the previous section, we adjust the harvesting rate? Can the population be saved? By what value should the rate be adjusted? The answers, at least for the model, lie in the realm of mathematics.

First let's find out what the level of the population is after 10 years. Using F3 - Trace will give us the answer as long as we have first set the initial condition in the Y= directory. See the screens that follow.



Thus, we need to find a value for h so that the unstable equilibrium is less than 2.2529399. Looking at the roots of the quadratic equation, and using the fact that a = 1 and b = 5, we have

$$\frac{5}{2} - \frac{\sqrt{25 - 20h}}{2} < 2.2529399$$

or

$$h < 1.2377922614$$

Contents

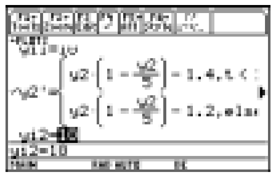
Obviously, a calculator was used in the preceding computations.

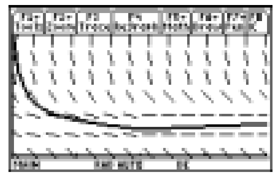
Let's assume that the harvesting rate is set at $h = 1.2$ when $t = 10$. In TI-89 or 92 syntax this is expressed as:

$$y' = \text{when}(t<10,\ y*(1-y/5) - 1.4,\ y*(1-y/5) - 1.2)$$

The numerical solver handles this type of equation as easily as it handles any of the other, less complicated expressions.

We store this new equation as y2 in the Y= editor and look at the flow field. We will also show the status of the population when $y(0) = 10$.





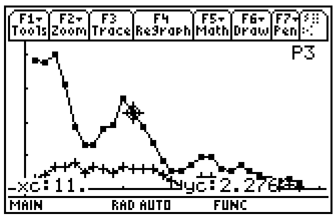
Modelling Real Data

The following data was gathered between the years of 1932 and 1959 in the Monterey Bay, California. It concerns the sardine population in the bay. The units are not individual sardines, but the total biomass of the population and the yearly harvest. We, of course, could convert these numbers into approximate populations by dividing the totals by the average weight of a sardine. This is not at all necessary.

Year	*Biomass Pop.*	*Biomass Catch*
1932	3.824	0.295
1933	3.764	0.387
1934	3.996	0.638
1935	3.136	0.632
1936	1.861	0.791
1937	1.330	0.498
1938	1.324	0.671
1939	1.772	0.583
1940	1.940	0.493
1941	2.709	0.680
1942	2.276	0.573
1943	1.849	0.579
1944	1.389	0.614
1945	0.835	0.440
1946	0.506	0.248
1947	0.524	0.130
1948	0.687	0.189
1949	0.958	0.339
1950	0.973	0.353

1951	0.570	0.145
1952	0.554	0.014
1953	0.425	0.018
1954	0.558	0.080
1955	0.425	0.078
1956	0.293	0.047
1957	0.212	0.032
1958	0.281	0.126
1959	0.190	0.059

If we enter this data using the Data/Matrix editor of the TI-89 and plot the result, we have a graphical display of the data.



Obviously this data does not have the shape of the solutions to the differential equations given in the previous section. Between 1935 and 1941 there is a dip in the sardine population that is not really related to the catch since the population recovers while the catch remains reasonably stable. The remainder of the graph appears as if it may be explained by the differential equation model with some modifications. The main modification is that the rate of change of the harvest is not constant. We will use the following model for the years 1935 – 1959

$$\textit{rate of change} = a*pop*(1-b*pop) - c*catch$$

This does not look different from our previous model until we realize that the catch is not constant. This model is more complicated than our previous model and **can not be solved analytically.** The reason is that we do not have a representative function for the catch and none of the regression techniques will be particularly useful for finding an approximating function.

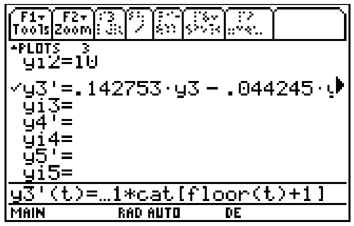
Our first job is to approximate the parameters, *a*, *b*, and *c*. we will use the data between 1935 and 1942 to arrive at a least squares approximation to these parameters. The basic equation that we will use is:

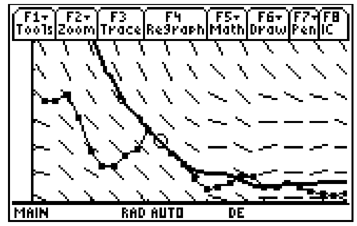
$$\begin{aligned} pop(t+1) - pop(t) &= a * pop(t) * (1 - b * pop(t)) - c * catch(t) \\ &= a * pop(t) - a * b * pop(t)^2 - c * catch(t) \\ &= a * pop(t) - \beta * pop(t)^2 - c * catch(t) \end{aligned}$$

Using the Data/Matrix editor and several matrix commands we can "solve" this over determined system for a least squares approximation for *a*, β, and c. We use the QR-Method, but that is a subject for another, more advanced, workshop. The values for the parameters, given our data, are

$a = 0.142753$ $\beta = 0.044245$ $c = 0.851010$

Let's see how it all worked out.





Obviously, when the model is out side of the range we used, $10 \le t \le 27$, The curves do not agree. Within the range, we seem to have a fair agreement (not perfect). The model also tends to be more optimistic towards the end of the period due to the small catch during those years.