6.8 Logistic Models

Logistic models describe limited growth, as opposed to unlimited, exponential growth. Examples of such limited growth are populations with a constraining factor (food, space, etc.), spread of diseases, and cumulative quantities. We have already seen examples of logistic models, such as the number of countries issuing postage stamps. We will introduce logistic models through the example of the spread of a disease (flu, AIDS, etc.) in a population into which one infected individual has been introduced. Before deriving the model equations, we will simulate the spread of the disease with the **Disease Simulation Game**.
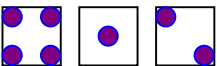
Imagine you are throwing a party where people interact with each other randomly. None of the guests knows who is healthy or infected. Every time a healthy person interacts with an infected person, there is a chance that the disease is spread.

Instructions on how to play the game:
Each student in the class gets the handout for the Disease Simulation Game (see last page of this section) and a die. Alternatively, the die rolls can be "created" using the TI-89 function rand(6), which simulates the random selection of a number from the set {1, 2, 3, 4, 5, 6} (see Assignment #2 in lesson T2).

Step 1: Creating ID numbers
Roll your die three times and write down the numbers you rolled to form your 3 digit ID number.
Example: If you rolled , then your ID is 412.

Step 2: Recording of ID numbers
The ID numbers of all students are recorded on the board, grouped according to the first digit. If the same ID number shows up more then once, the person with the duplicate ID rolls again until an ID is created that has not been used already.

Step 3: Playing a round of the game
Now the party begins. The instructor signals the beginning of the round (each long enough for 2 - 4 encounters, depending on the total number of students present). In each round, walk up to a person and exchange your ID number (and name if you wish) with that person. Mark your partner's ID number on the handout in the column corresponding to the current round. On a flat surface, each of the two partners rolls his/her die. If the **sum** of the two rolls is **less than or equal to 5**, then you've had a risky encounter (i.e., if one of the two partners has the disease, then the other one will

be infected). If you've had a risky encounter, circle your partner's ID on your sheet. Here is an example of how your sheet may look after round 2:

Round 1	Round 2	Round 3	Round 4	Round 5
413 (235) 152	(123) 112 345 (521)			

This would indicate that you had a risky encounter in round 1 with person #235, and risky encounters with persons #123 and #521 in round 2. Continue to meet additional people until the instructor signals the end of the round (and the beginning of the next round). Once the signal is given, finish any encounters that are already in progress. Play four to five rounds of the game.

Step 4: Determining the initially infected individual

The instructor rolls a die until it matches one of the first digits of the ID numbers listed on the board. Then the die is rolled until the number matches one of the second digits of those ID numbers that start with the first digit rolled. Finally, the die is rolled until it shows a number that forms a listed ID together with the two previously rolled digits. (Note: Any other system that randomly selects one ID can be substituted.) Circle the ID of the infected individual on the board.

Step 5: Tracing the spread of the disease

<u>Round 1</u>: Each person checks whether (s)he had contact with the infected individual in the first round. If the ID of the infected individual is circled in the column for round 1, then the person had a risky encounter with an infected individual and is now also infected. ID numbers of all such newly infected individuals are circled on the board. Record the number of newly infected players, as well as the total number of infected individuals for round 1 on your handout.

<u>Round 2 - 5</u>: For each round, anybody that is still healthy checks whether (s)he had a <u>risky</u> encounter with any person whose ID is circled on the board. If so, that person is recorded as infected and his/her ID is circled on the board. The number of newly infected and the total number of infected individuals are recorded for each round. Once you become one of the infected individuals, you no longer have to check your interactions.

Step 6: Graph of the data

Now you can graph the total number of infected people for each round on your handout.

Here is an example of the results of this game played by 30 mathematics professors at the 1998 ICTCM conference in New Orleans:

Round #	Number of Infected Individuals
0	1
1	3
2	5
3	13
4	23
5	29
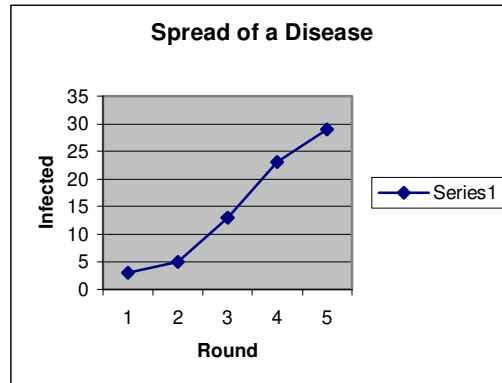
This graph looks very much like a logistic function. Initially, the number of infected individuals roughly doubles in each round, but then the number of infected people levels off (as we knew it would have to--after all there were only 30 people at the party). We will now derive the logistic model analytically and use the simulation to help us formulate the underlying mechanism.

First, let's define the variables for the model:

$$n \quad = \quad \text{number of rounds}$$
$$I(n) \quad = \quad \text{number of infected people after round } n.$$

Again using the paradigm new = old + change, we get

$$I(n) = I(n-1) + \text{newly infected people}.$$

What are the assumptions about how people get infected? Well, in order to have a newly infected person, a healthy individual has to be in contact with an infected individual. How many such encounters are possible? Let's look at an example:

If there are 3 healthy (H_1, H_2, H_3) and 2 infected (I_1, I_2) people, how many encounters of a healthy and an infected individual are possible?

Answer: We can pair each of the three healthy individuals with each of the two infected individuals, so there are $3 \cdot 2 = 6$ possible pairings. Below is a listing, as well as a visualization using a tree diagram.

Possible pairings = $\{(H_1, I_1) , (H_1, I_2), (H_2 , I_1), (H_2 , I_2), (H_3, I_1), (H_3, I_2)\}$
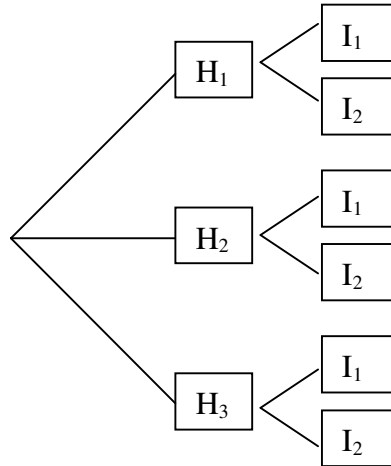
In general, the number of possible encounters that can lead to a newly infected person is given as the product of the number of healthy people and the number of infected individuals. However, not each encounter leads to an actual spread of the disease. In our game we simulated this by rolling the dice, and only a certain percentage (those where the sum of the dice was five or less) resulted in a spread of the infection. We can actually compute the probability for a risky encounter in our game. To do so, we need to find out the probability that the sum of the dice is five or less. Below is a chart of all the possible pairings of two dice. If we assume that the dice are all balanced and that each outcome is equally likely, then the probability of a risky encounter is given by

$$\frac{\text{number of dice rolls whose sum is} \leq 5}{\text{number of all possible dice rolls}}.$$

The dice rolls whose sum is less than or equal to five are highlighted in the table of possible pairings below.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

This leads to

$$\frac{\text{number of dice rolls whose sum is} \le 5}{\text{number of all possible dice rolls}} = \frac{10}{36} = .2778 \approx 28\%$$

i.e., the risk factor for being infected is 28% in this particular game. One final consideration that enters into the model is the fact that not all possible encounters take place in each round. For the mathematicians at ICTCM, each person met an average of 5 people per round (due to the time limit). Using the same idea as above, we can compute the probability for an encounter as the number of actual encounters over the number of possible encounters (this assumes, however, that all the encounters are equally likely). The total number of possible encounters among 30 people is given by $30 \cdot 29 / 2 = 435$. (Each of the 30 persons can meet all of the 29 others; however, if I meet you, then you meet me, so we have counted our encounter twice. To fix this error, we divide by 2.) Thus, the encounter rate is given by $\frac{(30 \cdot 5)/2}{435} = 0.172$ (each of the 30 people meets 5 people; divide by 2 to avoid double counting). We can now derive the number of newly infected people as follows:

# of newly infected

= (encounter probability)·(risk factor)·(# of infected people)·(# of healthy people)

= constant·(# of infected people)·(# of healthy people).

If we let the total number of people at the party be denoted by *N*, then the number of healthy people at time *n*-1 is given by

# of healthy people = (total number of people) – (# of infected people) = *N* - *I*(*n*-1)

and the number of newly infected individuals can be expressed as

newly infected individuals = *c*·*I*(*n*-1)·(*N*- *I*(*n*-1)).

This leads to the following iterative model equation for the logistic model:

I(*n*) = I(*n*-1) + *c*·I(*n*-1)·(*N*- I(*n*-1))

Remarks: 1) To make predictions with this model, we need to know the parameters *N* and *c*.

2) The quantity *N* - *I*(*n*-1) is often called *unused growth potential*. This terminology comes primarily from population models, where *N* is called *carrying capacity*, i.e., the maximal population size due to food, resource, and environmental limitations. As this level corresponds to the largest possible value, it represents the level of the larger of the two horizontal asymptotes of a logistic function.

3) When $I(n\text{-}1)$ is small, then the logistic function behaves almost like an exponential function. This can be seen if we distribute the terms in the above iterative model equation:

$$\begin{aligned} I(n) &= I(n-1) + c \cdot N \cdot I(n-1) - c \cdot I(n-1)^2 \\ &\approx I(n-1) + c \cdot N \cdot I(n-1) = (1+cN) \cdot I(n-1) \end{aligned}$$

where we used the fact that the square of a small number ($I(n\text{-}1)$) is even smaller and can be neglected. The resulting iterative model equation has the form of an exponential model (the new output value is proportional to the old output value).

Let's see how the logistic model derived above compares to the data from New Orleans. We can compute model predictions by defining a recursive sequence, then look at the table of values. To set up the model equation, we use the parameter values

$N = 30$ and $c =$ (encounter probability)·(risk factor) = (0.172)·(0.28) = 0.048.

The iterative model equation is therefore given by

$$I(n) = I(n-1) + 0.048 \cdot I(n-1) \cdot (30 - I(n-1)).$$

1. Press $\boxed{\text{MODE}}$ and set **Graph** to **SEQUENCE→**. Press $\boxed{\text{ENTER}}$, then $\boxed{\bullet}\boxed{\text{Y=}}$, then $\boxed{\text{F1}}$ **8** $\boxed{\text{ENTER}}$. This clears the screen of any previously defined sequences. Now type
    $\boxed{\text{alpha}}$ [U] **1** $\boxed{(}$ $\boxed{\text{alpha}}$ [N] $\boxed{-}$ **1** $\boxed{)}$ $\boxed{+}$ $\boxed{.}$ **048** $\boxed{\times}$
    We can use the copy and paste feature to avoid retyping the expression u1(n-1). Press the $\boxed{\triangleleft}$ key repeatedly until the cursor has moved to the right of the closing parenthesis. While pressing down the $\boxed{\text{↑}}$ key, use the $\boxed{\triangleleft}$ key until the expression you want to copy is highlighted.
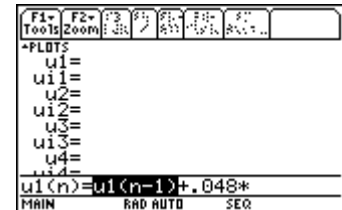
2. Press $\boxed{\bullet}\boxed{\text{COPY}}$ to copy this expression. Use $\boxed{\text{2nd}}\boxed{\triangleright}$ to move to the end of the entry line, then press $\boxed{\bullet}\boxed{\text{PASTE}}$. Now type
    $\boxed{\times}$ $\boxed{(}$ **30** $\boxed{-}$ $\boxed{\bullet}\boxed{\text{PASTE}}$ $\boxed{)}$ $\boxed{\text{ENTER}}$.
    Type **1** as the initial value and press $\boxed{\text{ENTER}}$.
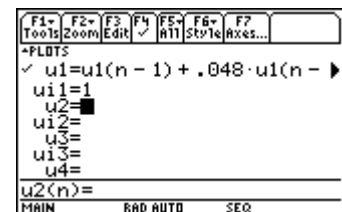
3. Press $\boxed{\bullet}\boxed{\text{TblSet}}$ and set **tblStart** to **0**, **Δtbl** to **1**. Make sure that **Graph ↔ Table** is set to OFF→ and **Independent** is set to AUTO→. Press $\boxed{\text{ENTER}}$, then $\boxed{\bullet}\boxed{\text{TABLE}}$ to see the values.

Below is a table comparing data and model predictions. Notice that since the model does not predict integer values, we need to round <u>down</u> the model predictions to the nearest integer.

| Round #	$I(n)$ (Data)	$I(n)$ (Model prediction)	Rounded Prediction
0	**1**	1	**1**
1	**3**	2.392	**2**
2	**5**	5.5618	**5**
3	**13**	12.086	**12**
4	**23**	22.478	**22**
5	**29**	30.594	**30**

Comparing the data with the (rounded) model predictions shows fairly good agreement, validating the logistic model.

Unlike in the case of a simulation, where we can compute the values for N and c, in a "real-life" situation we need to derive these values from the given data and its graph. While it is not easy to derive the value of c, we usually can determine the value of N, using our knowledge about the logistic function. Recall that a logistic function has two horizontal asymptotes, at level d (smallest possible value) and at level $a + d$ (largest possible value). Furthermore, a change in curvature occurs at (output) level $a/2 + d$.

So how do we determine the value of N, the largest possible value, from the data? There are two possibilities: Either we have enough data to see where the graph levels off eventually, or we only see the change of curvature, but not the leveling off. (If we do not even see the change in curvature, then we are dealing with an exponential function instead.) In the first case, we just read off the value of N from the graph and are done. In the second case, we need to use that $N = a + d$, i.e., we need to determine those two values from the data and the portion of the graph that is visible.

We start by determining the smallest possible value either from the graph or from the context. This will give us the value of d (level of the lower asymptote). Next we read off the output value at which the change in curvature occurs. This gives us an equation involving a and d, and we can substitute the value of d that we have found before. The resulting equation in the variable a can be solved quite easily. Here is how the procedure works in the case of the data from New Orleans, which is repeated below for easy reference.

Round #	Number of Infected Individuals
0	1
1	3
2	5
3	13
4	23
5	29
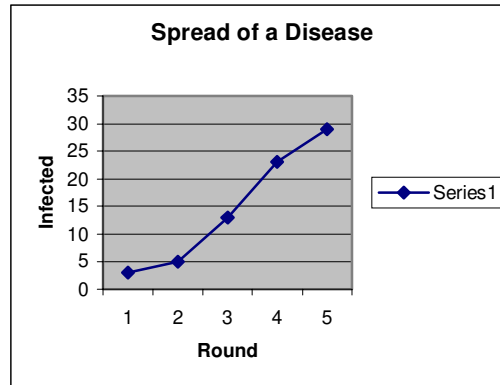
Looking at the graph, it is not quite clear where it may level off eventually. Thus, we start by determining the smallest value. We know that initially there was one infected person, i.e., $d = 1$. Next, we look at the change in curvature. It looks as if this change occurs around 16 or 17, so we will do our computations with the average of the two values, 16.5. This leads to the equation

$$\text{change in curvature} = a/2 + d = 16.5$$

Now substitute $d = 1$ into this equation and solve for a:

$a/2 + 1$	$= 16.5$	substitution
$a/2$	$= 15.5$	subtract 1 from both sides
a	$= 31$	multiply each side by 2.

Thus, we would predict the value of N to be $a + d = 31 + 1 = 32$. (Note that in this specific example, we know the value of N to be 30, the total number of people present, hence the largest number of people that could ever become infected. However, in "real" applications, this value is not known beforehand.)

Let's now fit a logistic function to the New Orleans data and see whether the values of a and d of the fitted function are similar to what we have derived by looking at the graph and using basic methods for solving an equation for its unknown.

1. Press APPS **6 3** to open a new data set. Set **Type** to **Data→** and **Folder** to **main→**. Name the data **disease** by typing
    2nd a-lock [D][I][S][E][A][S][E]
    and press ENTER.

2. Press $\boxed{\text{ENTER}}$, then $\boxed{\text{alpha}}$ to turn off the alpha lock. Enter the values for the round # in column c1, and the values for $I(n)$, the number of infected persons at time n, in column c2.

3. Press $\boxed{\text{F2}}$ to set up the data plot, then press $\boxed{\text{F1}}$. In the window, set **Plot Type** to **Scatter→** and **Mark** to **Square→**. For x type $\boxed{\text{alpha}}$ [C] **1**, for y type $\boxed{\text{alpha}}$ [C] **2** and for **Freq and Categories** select **NO→**. Press $\boxed{\text{ENTER}}$.

4. Press $\boxed{\text{ENTER}}$, then $\boxed{\text{MODE}}$ and set **Graph** to **FUNCTION→**. Press $\boxed{\text{ENTER}}$ followed by $\boxed{\bullet}$[WINDOW] and set the parameter values as follows:

   xmin = 0   xmax = 6   xscl = 1
   ymin = 0   ymax = 30   yscl = 2   xres = 2.
   Press $\boxed{\text{ENTER}}$.

5. Press $\boxed{\bullet}$[Y=] to deselect any other marked functions. One way to do this is to press $\boxed{\text{F5}}$ **1** (deselects everything), then highlight **Plot1** and to press $\boxed{\text{F4}}$. Then press $\boxed{\bullet}$[GRAPH] to see a graph of the data.
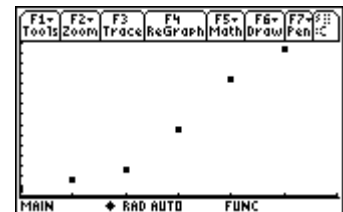
6. Press $\boxed{\text{APPS}}$ **6 1** to return to the data set. Select $\boxed{\text{F5}}$ and set **Calculation Type** to **C:Logistic→**. For x type $\boxed{\text{alpha}}$ [C] **1**, for y type $\boxed{\text{alpha}}$ [C] **2** and select an unused function in which to store the fitted function, for example **y1(x)**. Finally, set **Freq and Categories** to **NO→**.

7. Press $\boxed{\text{ENTER}}$. The resulting screen indicates that an logistic function of the form $y = \dfrac{a}{1 + b \cdot e^{c \cdot x}} + d$ has been fitted, where

   $a = 30.774343$, $b = 98.20466$, $c = -1.37127$, $d = 1.1.21891$.

8. Press $\boxed{\text{ENTER}}$, then $\boxed{\bullet}$[GRAPH] to see a graph of the data together with the fitted function.
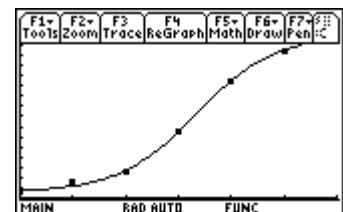
The fitted function looks quite good and the values for a and d are close to the ones we derived from looking at the graph and the context ($a = 31$ and $d = 1$). The fitted function would predict a carrying capacity of $N = a + d = 30.774343 + 1.121891 = 31.896234 \approx 31$ (rounded down), which is actually a better prediction (as we know that the "answer" in our special example is given by $N = 30$). The fact that the model predicted a carrying capacity of approximately 31, with the limited data set of 5 points, is quite good. If the simulation had continued, the data values would eventually settle down at 30, and the model would predict a carrying capacity of 30 as well.

How Diseases Spread (with Dice!)

Your ID Number _____

Round 1	Round 2	Round 3	Round 4	Round 5

Round #	# of newly infected people	Total # of infected people
0	1	1
1		
2		
3		
4		
5		

Total Number of Infected People

0   1   2   3   4   5

Rounds

Chapter Review

<u>Key Terms</u>

discrete dynamical system	linear system	Cob-web diagram
iterative model equation	geometric sum	birth rate
explicit model solution	compounded interest	death rate
half-life	stable equilibrium	growth rate
proportional	unstable equilibrium	growth factor
equilibrium	semi-stable equilibrium	logistic model
first-order system	neutral equilibrium	unused growth potential
second-order system	fixed point	carrying capacity

<u>Short Answers</u>

1. What is a discrete dynamical system?

2. Describe the difference between the iterative model equation and the explicit model solution.

3. Describe the method by which the explicit model solution is derived from the iterative model equation.

4. How do you determine the order of a DDS?

5. Explain the terms "unused growth potential" and "carrying capacity" in the context of a population model.

6. Compare the formulas for half-life and the time it takes to double money. How do they differ and what do they have in common?

7. What is the difference between the two types of graphs for sequences (**Axes** set to **WEB→** versus **Axes** set to **TIMES→**)?

8. How do you determine whether a DDS is linear?

9. Describe how you can use data on the size of a population to derive its growth factor.

10. Describe how to find equilibrium values for a non-linear DDS (both analytically and graphically).

11. Describe the different types of population models and where they are applied.

True - False Questions

T F   1. If the output value is reduced by a fixed percentage per time unit, then the amount removed in each time unit is constant.

T F   2. In a linear first-order DDS, the values of $x(n)$ will increase or decrease without bound if $|a| < 1$.

T F   3. A second-order DDS always contains both $x(n\text{-}1)$ and $x(n\text{-}2)$ terms on the right hand side of the iterative model equation.

T F   4. $x(n) = x(n\text{-}1)\, x(n\text{-}2)$ is a linear DDS.

T F   5. A savings account which pays 3% interest compounded 4 times a year pays a total of $4 \cdot 3\% = 12\%$ interest per year.

T F   6. In a Cob-web diagram, the axes display the input and output values.

T F   7. The equilibrium value for a first-order linear DDS always exists.

T F   8. The growth rate and the growth factor are the same.

T F   9. An equilibrium value x is unstable if sequences that start nearby sometimes tend toward and sometimes tend away from the equilibrium value.

T F   10. For small input values, the logistic model behaves almost like an exponential model.

Fill in the Blanks

1. The paradigm _____ is used to set up the iterative model equation.

2. Two quantities x and y are proportional if there exists a constant k such that _____.

3. The Cob-web diagram visualizes the _____.

4. To make predictions for the population model using the explicit solution $P(n) = (1 + c)^n P(0)$ you need to know the _____ $P(0)$ and the _____ $(1+c)$.

5. The time it takes to remove half the original amount of drug is called _____.

6. A_____ model is characterized by limited growth.

7. If the growth factor is bigger than one, then the population _____.

8. A sum of the form $1 + a + a^2 + \cdots + a^n$ is called _____.

9. The growth rate is the difference between _____ and _____.

10. The iterative model equation for a logistic model is given by _____.

Facts (for first-order linear DDS)

Iterative model equation: $x(n) = a \cdot x(n\text{-}1) + b.$

Explicit solution: $x(n) = \begin{cases} a^n x(0) + b \dfrac{1 - a^n}{1 - a} & \text{if } a \neq 1 \\[2mm] a^n x(0) + b \cdot n & \text{if } a = 1 \end{cases}$

Equilibrium value: $x = \begin{cases} \dfrac{b}{1 - a} & \text{if } a \neq 1 \\[2mm] \text{any value} & \text{if } a = 1 \text{ and } b = 0 \end{cases}$

Stability of equilibrium: $x \text{ is } \begin{cases} \text{stable} & \text{if } |a| < 1 \text{ or } a = 1 \text{ and } b = 0 \\ \text{unstable} & \text{if } |a| > 1 \\ \text{neutral} & \text{if } a = -1 \end{cases}$