Contents

MATH 221 PROJECT

Functions and Mathematical Models

1. Purpose of this project

An understanding of calculus can only be achieved if you, the student, have a firm grasp on the concept of a function. Literally everything in this, and future, courses depends on it. Therefore, in this project we will explore what is meant by the term *function*, and see how functions arise naturally in a number of application areas. Further, we will think about what a "mathematical model" is, what they are used for, and the limits of their use.

2. Background on Functions

We recall from the class notes the following definition.

2.1. Definition. *A function f is a rule giving a value, denote $f(x)$, for each x. The set of x for which f is defined (i.e. for which the rule can be applied) is called the* ***domain*** *of f, and the set of all possible values, $f(x)$, is called the* ***range****. That is, the range of f is the set*

$$\{f(x) \mid x \text{ is in the domain of } f\}.$$

All of the functions encountered in math 221 will have a domain and range that are *subsets* of the real numbers. Note that the definition of a function is very broad and so particular functions can be defined in a number of different ways as the next few examples demonstrate.

2.2. Example. Consider a vendor selling pizza by the slice on State Street. Suppose you can purchase 1, 2, 3, or 4 slices of pizza at a cost of \$2, \$3.75, \$5.25, and \$6.50, respectively. The vendor does not take any other orders (that is, he will not allow you to purchase 5 slices since that is greedy).

If we denote by s the number of slices purchased, and by $p(s)$ the price of that order, in dollars, then p is a function with domain $\{1, 2, 3, 4\}$ and range $\{2, 3.75, 5.25, 6.50\}$. The rule defining p is then given by the following list

$$p(1) = \$2, \quad p(2) = \$3.75, \quad p(3) = \$5.25, \quad p(4) = \$6.50.$$

□

2.3. Example. A scientist at the Wisconsin Institute for Discovery believes that for every virus particle that infects a group of cells, one thousand new virus particles will be made. If we let v denote the number of virus particles that infect a group of cells, and by $f(v)$ the number of resulting virus particles, then $f(v)$ is a function with domain $\{0, 1, 2, \dots\}$ and range $\{0, 1000, 2000, \dots\}$ defined by the rule

$$f(v) = 1000 \cdot v.$$

Note that because the domain is no longer finite, it is no longer possible to give a list detailing the rule.□

2.4. Example. A car is moving away from you at a speed of 30 miles per hour. At time zero, the car was already 2 miles away. Let $t \geq 0$ denote time (in hours), and let $d(t)$ denote the distance between you and the car. Then $d(t)$ is a function with domain $[0, \infty)$ and range $[2, \infty)$ defined by the rule

$$d(t) = 2 + 30t$$

□

Of course, each of the functions above was derived from some (made up) "real world" application. In a math course we will typically consider functions with no overt connection to the "real world." For example, as in the class notes on page 9, we can define $f(x)$ to be the function defined piecewise via

$$f(x) = \begin{cases} 2x & \text{for } x < 0 \\ x^2 & \text{for } x \geq 0 \end{cases}.$$

Having such flexibility allows us to play with an infinite number of examples.

2.5. Functions and Mathematical models. A *mathematical model* is a mathematical description of some real world phenomenon. Usually the model is described via an equation or a system of equations. The purpose of a mathematical model is to gain insight into the phenomenon being studied and, hopefully, to make predictions.

2.6. Example. Consider a large manufacturer of Bucky Badger shirts. Every year, the company spends $200,000 on fixed operational costs (wages, electricity, mortgage, etc.). Further, every shirt made costs an addition $5. Let q denote the number of shirts produced by the company in a given year, and let $C(q)$ denote the cost to manufacture q shirts. Assume that if zero shirts are ever produced, they will fire all their employees and close their factory yielding a manufacturing cost of zero. □

3. Problems

1. What is the domain of the function C? What is the range?

2. A quantity of great interest to many manufacturers is the *marginal cost*, defined as the cost required to produce one addition unit. That is, the marginal cost at 5 units is the difference in the cost of 6 and 5 units. Define a function giving the marginal cost for our manufacturer. Be clear in specifying a domain, range, and rule.

3. Our manufacturer does not only care about costs, but also wants to know about revenue. Suppose they sell each unit for $13. Write down a function for revenue, $R(q)$, generated by selling the Bucky Badger shirts, being careful to specify an appropriate domain and range. How many shirts would need to be sold to make it worthwhile to manufacture them? Be explicit in how you solved this problem. A graph would certainly help.

4. Relative risk and the effects of exercise on rates of breast cancer

According to the Susan G. Komen foundation for the Cure[1], we define the *absolute risk* as a person's chance of developing a certain disease over a certain period of time. Next, a *relative risk* is calculated by comparing two absolute risks. The numerator (the top number in a fraction) is the absolute risk among those with the risk factor. The denominator (the bottom number) is the absolute risk among those without the risk factor. Hence,

$$\text{Relative risk} = \frac{\text{Absolute risk with factor}}{\text{Absolute risk without factor}}.$$

Note therefore, that if the relative risk of an activity is above 1.0, then the risk is higher for those people performing that activity. If, on the other hand, the relative risk is below 1.0, then the risk is lower for the portion of the population performing that activity is lower than the population at large.

[1] See, http://ww5.komen.org/BreastCancer/UnderstandingRisk.html

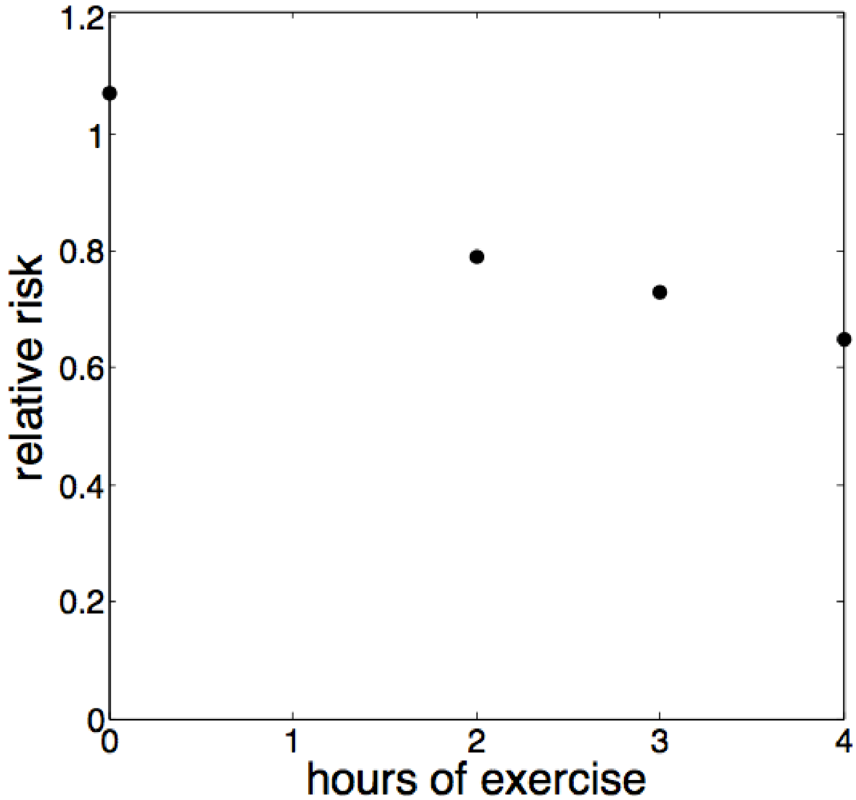


Figure 1. Relative risk of breast cancer vs. hours of exercise.

We present the following case study, that is based on a similar study performed in Duke University's Laboratory Calculus course. For reference, the data used in this study is taken from a study of women, age 50 - 64, that were residents of King County in northwestern Washington State.[2]

One risk factor considered in the King County study was exercise level. The researchers asked each participant in the study "During the 2-year period prior to (date of study), with what frequency did you do any strenuous physical activities, exercise, or sports?" The following data was collected:

Average number of exercise hours per week	0	2.0	3.0	4.0
Relative Risk	1.07	0.79	0.73	0.65

This data in the table above is plotted in Figure 1

[2] *Occurrence of Breast Cancer in Relation to Recreational Exercise in Women Age 50 - 64 Years*, Epidemiology, Nov. 1996, Vol. 7, No. 6, 598 - 604

5. Problems

1. **(a)** Draw a straight line in Figure 1 that fits the data well (do not use the statistical features of your calculator). Denote the relative risk by the variable R, and hours of exercise by the variable E. Determine the slope of the line you drew. Also determine the R intercept. What is the equation of the line?

(b) What does the R intercept tell you about exercise and the relative risk of breast cancer? is your intercept the same as your data point on the axis? Which do you think is more accurate? Why?

(c) The slope of your line should be negative. Describe what this means in terms of relative risk of breast cancer and exercise.

(d) What is the absolute value of the slope? Use this number and the negativity of the slope of the line to finish the following sentence: "If a woman increases her exercise by one hour per week, then her relative risk of breast cancer will be (approximately) …."

(e) Estimate the relative risk of breast cancer if a woman exercises for 3.3 hours per week.

(f) Find the E intercept (the place where the line intersects the horizontal axis), and denote it by M. Interpret M.

(g) For what values of E do you believe your linear model to be a decent approximation to the true relative risk?

Report Instructions

Using complete sentences (including punctuation!), write out your solutions to the exercises found in Example 2.6. Next, carefully write up your answers to the questions found in Example 4. You should include not only the formulas used, but also all the derivations of the formulas. General instructions:

(1) Do not turn in your first draft. Instead, solve all the problems first, and then write or type up a clean report.
(2) Only one report is required per group, but all members must take an active role in the solution and writing of the report. If a member does not participate, his or her name should not be included on the final report.

MATH 221 PROJECT

Models for the dynamics of populations

1. Purpose of this project

We will build mathematical models for the rates of change of different populations from first principle arguments. We will see that the same mathematical model can be used to describe such disparate things as populations of living organisms and chemical concentrations.

2. Malthusian Population Growth

Consider a population living in isolation with an abundance of resources. Let $P(t)$ denote the size of the population at some time, t. Note that $P(0)$ is the size of the population at time zero (where we can define "time zero" to be whenever we want: 0 A.D., January 1st 2012, yesterday, etc.). We now make the *assumption* that for any small $h > 0$, representing a small window of time, the growth of the population from time t to time $t + h$ should be proportional to both

(1) the size of the population, and
(2) the size of h.

Mathematically, we assume that

$$P(t+h) \approx P(t) + \lambda P(t)h, \tag{1}$$

for some $\lambda > 0$ (the constant of proportionality). Rearranging terms yields

$$\frac{P(t+h) - P(t)}{h} \approx \lambda P(t). \tag{2}$$

The term λ can often be approximated from experimental data (How? The project for Chapter 6 will explain this in detail.). Note that the right hand side of equation (2) does not depend upon h. Further, the above is assumed to hold for *all* $h > 0$, and so letting h get very small we see that the left hand side is well approximated by the derivative of P at t (see page 20 of the notes):

$$P'(t) = \lim_{h \to 0} \frac{P(t+h) - P(t)}{h}.$$

Thus, we are led to the conclusion that the rate of change of $P(t)$ equals $\lambda P(t)$ for some $\lambda > 0$. Mathematically, we have that P satisfies the *differential equation*

$$P'(t) = \lambda P(t).$$

Note that this conclusion was reached logically from our assumptions above, which we consider to be our "first principles" for this model. If you do not believe one of the assumptions holds for a particular population, then you can not trust the conclusion!

We will not consider how to solve differential equations here (you will see these again in Math 222), however we observe that the above equation makes a startling claim:

the population will increase without bound forever!

Let's see why by returning to equation (1). For ease, we suppose that $P(0) = 1$, $\lambda = 1$, and we take $h = 1/2$. Using these values, fill in the following chart:

t	$P_{1/2}(t)$
0	1.00
0.5	1.50
1.0	2.25
1.5	
2.0	
2.5	
3.0	

where we are writing $P_{1/2}(t)$ for the size of the population since we are explicitly choosing an h of size $1/2$. Note that the population is not just growing, but the *rate* at which the population is growing is even increasing.

3. Problems

1. Make a similar chart as that above with $h = 1/4$. That is, calculate $P_{1/4}(t)$ for $t \in \{0, 0.25, 0.5, 0.75, \ldots, 3.0\}$. How do the solutions $P_{1/4}(t)$ and $P_{1/2}(t)$ compare for

$$t \in \{0.5, 1, 1.5, 2, 2.5, 3.0\}?$$

Why do you think this is happening?

4. The Logistic Growth Model

The fact that the previous model predicts continued growth, without bound, is a serious shortcoming for many reasonable situations. For example, if there is a limited amount of natural resources, unsustained growth is clearly not realistic. The following twist on the previous model fixes the problem.

We still consider a small time window $[t, t+h]$ and ask how the population changes over that time period. Now we make the following assumptions:

(1) There is a "carrying capacity" $K > 0$ such that if $X(t) > K$, the population should decrease and if $X(t) < K$, then the population should increase.
(2) The further from K the population is, the stronger the gain/decrease.

One such model satisfying these reasonable ideas is

$$X(t+h) \approx X(t) + \lambda X(t)\left(1 - \frac{X(t)}{K}\right)h. \tag{3}$$

Similarly to the Malthusian model, the associated *differential equation* satisfied by X is

$$X'(t) = \lambda X(t)\left(1 - \frac{X(t)}{K}\right). \tag{4}$$

The model described here is called the *logistical growth model.*

Returning to equation (3), complete the following charts for the different $X(0)$ assuming that $\lambda = 1$, $K = 4$, and $h = 0.5$.

t	$X(t)$
0	1.00
0.5	1.375
1.0	1.826
1.5	
2.0	
2.5	
3.0	

t	$X(t)$
0	7.00
0.5	4.375
1.0	4.170
1.5	
2.0	
2.5	
3.0	

What do you think each $X(t)$ converge to for larger and larger t?

1. What do you think each $X(t)$ converge to for larger and larger t? Why?

2. More generally, consider $X(t)$ that satisfies (4) for some $\lambda > 0$ and $K > 0$. What is the sign of $X'(t)$ when $X(t) > K$? What about when $X(t) < K$? Using these facts, argue what will happen to $X(t)$ as t gets very large. Try to be persuasive. (Note: proving this predicted behavior is a subject for a future mathematic course on *dynamical systems*.)

3. Check that that if X satisfies (3), then it satisfies the assumptions (a) and (b) above.

4. If X satisfies (3) for all $h > 0$, why should it satisfy the differential equation (4)?

5. Chemistry

We now change topics and consider the chemical reaction

$$A + B \to 2A.$$

Compare this with the chemical reaction found on page 21 of the notes. Note that every time the above chemical reaction occurs, we gain a molecule of A and lose a molecule of B. This implies that the total number of molecules, $A + B$, is *conserved*. Thus, if $[A](t)$ and $[B](t)$ represent the concentration of A and B at time t, then

$$[A](t) + [B](t) = M$$

implying

$$[B](t) = M - [A](t). \tag{5}$$

We now make the following assumptions, which are our *first principles*. The number of times the reaction happens in the the time window $[t, t + h]$ is proportional to

(1) the product of $[A]$ and $[B]$, and
(2) the size of h.

6. Problems

1. Use the first principles above to find a difference equation, similar to (3), for $[A](t)$. Note that, due to (5), $[B](t)$ should not appear in the equation.

$$[A](t + h) = [A](t) +$$

2. Using your result from Exercise 6.1, derive a differential equation for $[A](t)$:

$$\frac{d[A](t)}{dt} =$$

Compare the differential equation to (4).

Report Instructions

Using complete sentences, write a few paragraphs summarizing the project. Next, write up your solutions to all of the exercises. Be sure to include all pertinent information from the project itself. That is, a reader should be able to sit down with only your report and be able to understand what you are doing, and the conclusions you have drawn.

MATH 221 PROJECT

Limits

1. Purpose of this project

We will explore how limits can be used to solve interesting problems, some of which were "paradoxes" for thousands of years before calculus came to the rescue. In the next chapter of the notes, we will see how limits can be used to formally define a notion of a "derivative." In this project, we will use the limit for something else: trying to figure out distances traveled and areas under curves (which will later be called an *integral*). This project is theoretical in nature. However, students who put time into trying to fully understand the concepts involved will surely be rewarded for their efforts!

2. Overview

While the concept of the "limit" may seem hard to grasp, it is not an exaggeration to say that it is *the* conceptual breakthrough that allowed for calculus to be developed (and, hence, all the nice things that followed: air travel, mp3 players, etc.). As we have already seen, the limit is the necessary concept required to define a derivative. However, there is another object in calculus, the *integral*, that will be defined in the coming chapters and that relies on the concept of a limit as well. This project will include our first computation of an integral, though we will not formally make this connection until later in the semester. We will also point out how our computation of an integral is related to an ancient paradox.

3. A (seeming?) paradox

The following is a version of a paradox usually attributed to the Greek philosopher Zeno[3]:

Imagine that a person is walking towards a wall that is currently one mile away. It seems that this person should have no trouble walking to the wall. However, we will consider a different way to think about how much distance this person must travel to reach the wall:

(1) The person must walk half to the wall, yielding a total distance of 1/2 mile. This leaves 1/2 mile to go.
(2) The person must walk one half of the remaining 1/2 mile. This leaves 1/4 mile to go.
(3) The person must walk one half of the remaining 1/4 mile. This leaves 1/8 mile to go.

$$\vdots$$

Therefore, we have broken the interval of one mile down into an infinite number of finite pieces of length

$$1/2, \quad 1/4, \quad 1/8, \quad 1/16, \quad etc.,$$

and so the total distance that must be traveled must be infinity! Hence, *the person should never reach the wall!* Another way to put this is that the requirement of walking a mile has been broken into an infinite number of tasks (walking half of the remaining distance over and over again), and this should be impossible.

[3]See, for example, http://en.wikipedia.org/wiki/Zeno's_paradoxes.

3.1. A resolution. If the above argument does not sit well with you, then you have good intuition. Our resolution of the "paradox" rests on limits at infinity (page 29 of the notes). We will let $f(n)$ denote the distance traveled by our wanderer after completion of the first n "tasks" as described in the section above. Thus, we see

$$\begin{aligned} f(1) &= \frac{1}{2} \\ f(2) &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \\ f(3) &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{3}{4} + \frac{1}{8} = \frac{7}{8} \\ &\vdots \\ f(n) &= \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^n} = \frac{2^n - 1}{2^n} = 1 - \frac{1}{2^n}. \end{aligned}$$

Therefore, using the Limit Property (P4) on page 39 of the text, with the conclusion of Example 5.5 on page 37, we have that

$$\lim_{n\to\infty} f(n) = \lim_{n\to\infty} \left(1 - \frac{1}{2^n}\right) = 1 - \lim_{n\to\infty} \frac{1}{2^n} = 1.$$

Thus, even with the perspective of breaking the interval up into an infinite number of intervals, we arrive at the (reasonable) conclusion that the total distance traveled is one mile.

4. Area under a triangle

We now turn to a seemingly disparate topic: what is the area of a triangle? More specifically, consider the graph of the function $f(x) = x$ on the interval $[0, 1]$. The space between the x-axis and $f(x)$ for $x \in [0, 1]$ makes a triangle and we want to know the area. We will solve this problem in two ways. The first is using simple geometry, and will be familiar to everyone. The second uses limits and is a precursor to the much deeper idea of *integration* that you will be exposed to later in the course.

Solution 1. The triangle has height equal to one, and a base of size one. Hence, the area of the triangle is

$$\frac{1}{2} \cdot \text{base} \times \text{height} = \frac{1}{2}.$$

Solution 2. Limits commonly arise when we wish to make an approximation precise. This is what we will do here. More precisely, we will make a series of approximations, with the nth approximation denote by $A(n)$. We will only assume that we know how to compute the area of rectangle.

For $n \in \{1, 2, 3, \dots\}$, we will break the interval $[0, 1]$ up into n equally spaced intervals. That is,

(1) For $n = 1$, there is only one interval, $[0, 1]$.
(2) For $n = 2$, there are two intervals, $[0, 1/2]$ and $[1/2, 1]$.
(3) For $n = 3$, there are three intervals, $[0, 1/3]$, $[1/3, 2/3]$ and $[2/3, 1]$.
(4) The n intervals for an arbitrary integer n are

$$[0, 1/n], [1/n, 2/n], \dots, [(n-1)/n, 1].$$

Now, we will approximate the area of the space between the graph of $f(x)$ and the x-axis on each subinterval,

$$[i/n, (i+1)/n],$$

by the area of a rectangle, which we know how to compute. Since $f(x) = x$ for all x, we know that

$$f(i/n) = i/n$$

and we can approximate the area of the ith subinterval by

$$f(i/n) \times \frac{1}{n} = \frac{i}{n}\frac{1}{n} = \frac{i}{n^2}.$$

Adding up all these areas then yields the approximation (note that the first interval is always being evaluated at the left endpoint, or zero)

$$A(n) = \sum_{i=0}^{n-1} \frac{i}{n^2} = 0 \times \frac{1}{n^2} + 1 \times \frac{1}{n^2} + 2 \times \frac{1}{n^2} + \cdots + (n-1)\frac{1}{n^2}.$$

Note that n is fixed in the above sum!

5. Problems

1. Draw a detailed picture of the preceding argument. More specifically, for $n = 3, 4$, and 5,

(a) Plot $f(x) = x$ on the interval $[0, 1]$.

(b) Break the interval up into n evenly spaced subintervals.

(c) Show the area of the rectangles being computed.

(d) Compute the approximate area for each $n \in \{3, 4, 5\}$.

2. Let m be an arbitrary integer. We will compute

$$1 + 2 + \cdots + m.$$

To do so, let $S(m)$ be the (unknown) sum and add the following vertically:

$$\begin{array}{rl} & 1 \quad + \quad 2 \quad + \cdots + (m-1) + \quad m \quad = S(m) \\ & m \quad + (m-1) + \cdots + \quad 2 \quad + \quad 1 \quad = S(m) \\ \hline \implies & (m+1) + (m+1) + \cdots + (m+1) + (m+1) = 2S(m) \end{array}$$

Conclude that

$$S(m) = \frac{m(m+1)}{2}.$$

3. Show that

$$A(n) = \frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n}.$$

where $A(n)$ is the approximate to the area of the triangle after breaking $[0, 1]$ into n subintervals. Conclude that

$$\lim_{n\to\infty} A(n) = \frac{1}{2}.$$

Report Instructions

Write a few paragraphs describing Zeno's paradox and its resolution. In particular, note that the results from the class notes (page 29) allowed us to conclude that

$$\lim_{n\to\infty} \frac{1}{n} = 0,$$

whereas we used that

$$\lim_{n\to\infty} \frac{1}{2^n} = 0.$$

Why is this okay to do? Explain.

Write a few paragraphs describing how to compute the area of a triangle. Be sure to include solutions to all of the exercises. Note especially that the second method of computing the area is much harder than the first. However, what if the function we were approximating, f, where not a straight line, but had a "curve" to it (like a sine function)? Say a few words on how one may hope to find the area between its plot and the x-axis. Note that you *could not* simply resort to using a straight geometric argument like we could for the triangle. We will formalize this method (and call it the *definite integral*) in later chapters.

MATH 221 PROJECT

Derivatives

(and the important topic of ants walking on ropes)

Purpose of this project

In this project we will explore the following problem, which we will solve using the rules of differentiation (and a little integration, which will come later in the class).

Problem statement: [4] Suppose there is a rope that is very pliable stretched to be of size 1 meter. Suppose that one end of the rope is pinned down and an ant is placed at that end. At that moment, the ant starts walking at a rate of 1 meter per hour towards the other end. However, at the same time we begin stretching the rope (away from the pinned down side) so that the end of the rope is moving at speed 10 meters per second. Will the ant ever reach the end of the rope?

While this version of the problem may seem strange, it actually has relevance to a very reasonable question coming from cosmology: can the light from a distant galaxy ever reach us if the universe is expanding? At issue is when a galaxy is so far away that the relative speed between our galaxy and the other is greater than the speed of light due to the expansion of the universe. In this scenario, it seems that the light from that far flung galaxy should never reach us. However, in an analogous manner as the ant on the rubber rope, we can be sure it does without having to resort to a breakdown in the laws of physics.

1. Solution

We will solve the problem with the numbers given in the statement of the problem. The answer is "yes, the ant does reach the end." Surprisingly, the fact that the ant eventually reaches the other end does not depend upon the specific numbers given. Thus, even if the ant were walking at 0.0001 meters per second, and you were stretching this rope near the speed of light, the ant will still eventually reach the end. Of course, the *time* at which the ant reaches the end depends upon the rates chosen, so it would be more accurate to say that an immortal ant would eventually reach the end.

Any good solution to a hard word problem requires good notation! So, we begin with the following.

(1) Let $D(t)$ denote the distance of the ant from its starting point at time t. Thus, for example, $D(0) = 0$.

(2) Note that because the end of the rope is moving at a constant speed of 10 meters per second, and because the end of the rope as located one meter from the pinned down edge at time zero, the position of the end of the rope at time t is

$$1 + 10t.$$

(3) Let $P(t)$ denote the proportion of the rope that the ant has already traversed. Note that until the ant reaches the end, $P(t)$ is always between 0 and 1. For example, if $P(t) = 0$, then the ant is at the beginning of the rope (which is true when $t = 0$), and if $P(t) = 1$, then the ant has reached

[4] See also the wikipedia page on this problem at http://en.wikipedia.org/wiki/Ant_on_a_rubber_rope

the end. Also note that we can combine our knowledge of $D(t)$ and the position of the end of the rope to conclude that

$$P(t) = \frac{D(t)}{1+10t}.$$

With the above notation in hand, we can start thinking about how to solve the problem. It should be clear that we hope to show that

$$P(t^*) = 1,$$

for some finite t^*. However, to get there, we need to first consider $D(t)$.

2. The function $D(t)$

We need to understand how the position of the ant, $D(t)$, is changing with respect to time. The difficult issue is that the ant is moving due to two things

(1) Its own movement relative to the rope.
(2) The movement due to the pulling of the rope.

Whenever we have more than one action forcing something else to change we should try to break up the overall behavior into its components. Therefore, we should expect to have

$$D(t+h) = D(t) + \text{“movement due to ant only”} + \text{“movement due to rope”}.$$

2.1. Movement due to ant. Since the ant is moving at a speed of 1 m/h, and the amount of time that passes is h hours (which we think of as very small, of course), we have

$$D(t+h) = D(t) + 1 \cdot h + \text{“movement due to rope”}.$$

2.2. Movement due to rope. The movement due to the rope seems tricker. However, thinking along the following lines will save the day:

(1) We need to understand how fast a particular point on the rope is moving.
(2) We should recognize that for each point on a rope, it's relative position on the rope (i.e. the proportion of the rope behind (or in front of) it) should not change with time.

Thus, if $x \in [0,1]$ was a point on the rope at time zero, and $X(t)$ denotes its position at time t, then

(6) $$\frac{x}{1} = \frac{X(t)}{1+10t},$$

for all t. Since it holds for all t, we can conclude that

(7) $$\frac{X(t+h)}{1+10(t+h)} = \frac{X(t)}{1+10t},$$

since both sides equal the same thing (namely, $x/1$).

3. Problem

1. Explain why equation (6) should hold for *all* $t \geq 0$. Show that equation (7) implies that

(8) $$X(t+h) - X(t) = X(t)\frac{10h}{1+10t}.$$

3.1. Returning to $D(t)$. We return to our analysis of $D(t)$ armed with more knowledge. Since the position of the ant at time t is $D(t)$, equation (8) tells us that the total movement of the ant due to the stretching rope over a time period of size h is approximately

$$D(t)\frac{10h}{1+10t}.$$

Combining all of the above, we have

$$\begin{aligned} D(t+h) &= D(t) + \text{“movement due to ant only”} + \text{“movement due to rope”} \\ &= D(t) + h + D(t)\frac{10h}{1+10t}. \end{aligned}$$

Rearranging terms and dividing by h gives

$$\frac{D(t+h)-D(t)}{h} = 1 + D(t)\frac{10}{1+10t}.$$

Of course, we know that as $h \to 0$, the right hand side becomes $D'(t)$. Thus, we can conclude that

$$D'(t) = 1 + D(t)\frac{10}{1+10t}, \tag{9}$$

where we also know that $D(0) = 0$. The above equation tells us that the function D has the following properties:

(1) It has a value of zero at time zero (we already knew that, of course!).
(2) Its derivative is equal to itself multiplied by

$$\frac{10}{1+10t},$$

plus one.

4. So what?!?! – Change perspective.

Unfortunately, equation (9) is a differential equation, and is beyond the scope of this class to solve. That is, at this point in your mathematical education it is not clear at all that there should even be a function satisfying such an odd equation.[5] Therefore, this seems like a dead-end.

However, at this point, we can remember that it was not $D(t)$ that we were after anyways. Instead, it was the function giving the *proportion* of the rope already traversed by time t,

$$P(t) = \frac{D(t)}{1+10t}.$$

5. Problem

1. Show that

$$P'(t) = \frac{D'(t)}{1+10t} - \frac{10D(t)}{(1+10t)^2}.$$

Then use (9) to conclude that

$$P'(t) = \frac{1}{1+10t}. \tag{10}$$

[5] There is one. However, you will need more math courses to be able to solve these types of equations!

It is important to note that equation (10) is much simpler than equation (9) since the right hand side does not depend on anything except t. Thus, equation (10) can be read in the following manner:

"Find a function $P(t)$ whose derivative is $\dfrac{1}{1+10t}$."

Note also that we require $P(0) = 0$. Finding such a P is actually the exact *opposite* of what you have learned to do so far, which is how to *take* a derivative.

6. What to do?

Later in the course, you will find an explicit representation for $P(t)$ (if you must know now, the solution is $P(t) = (1/10)\ln(1 + 10t)$, where "ln" is the natural logarithm). For now, it is sufficient to know that there is a solution, and it yields a t^* at which

$$P(t^*) = 1.$$

For example, the value of t^* for our specific choice of constants is

$$t^* = \frac{e^{10} - 1}{10} \approx 2,202.5.$$

7. Problems

1. Redo the analysis of the project except assume that the ant is walking at a speed of s meters per hour (as opposed to one meter per hour) and that the end of the rope is being pulled at a rate of r meters per hour. Of course, assume that $s > 0$ and $r > 0$. You should conclude that in this (general) case

$$P'(t) = \frac{s}{1 + rt}.$$

2. Assume now that the ant is walking at 1 meter per hour and the rope is being pulled at a rate of 2 meters per hour. Using the result from the previous exercise, argue that for small h,

$$P(t + h) \approx P(t) + \frac{1}{1 + 2t}h. \tag{11}$$

Using this approximation for P, fill in the following tables to approximate t^* for which $P(t^*) = 1$:

t	$P(t)$ ($h = \frac{1}{2}$)
0.0	0.00
0.5	0.50
1.0	0.75
1.5	0.92
2.0	1.04
2.5	1.14

t	$P(t)$ ($h = \frac{1}{3}$)
0.00	0.00
0.333	0.33
0.666	0.53
1.000	
1.333	
1.666	
2.000	
2.666	
3.000	
3.333	

t	$P(t)$ ($h = \frac{1}{4}$)
0.00	0.00
0.25	0.25
0.50	0.42
0.75	
1.00	
1.25	
1.50	
1.75	
2.00	
2.25	
2.50	
2.75	
3.00	
3.25	
3.50	

For example, the first chart giving the computation with $h = 1/2$ yields an approximation

$$t^* = 2.$$

Report Instructions

Using complete sentences, write out solutions to all of the exercises.

Determining functional form from data

1. Purpose of this project

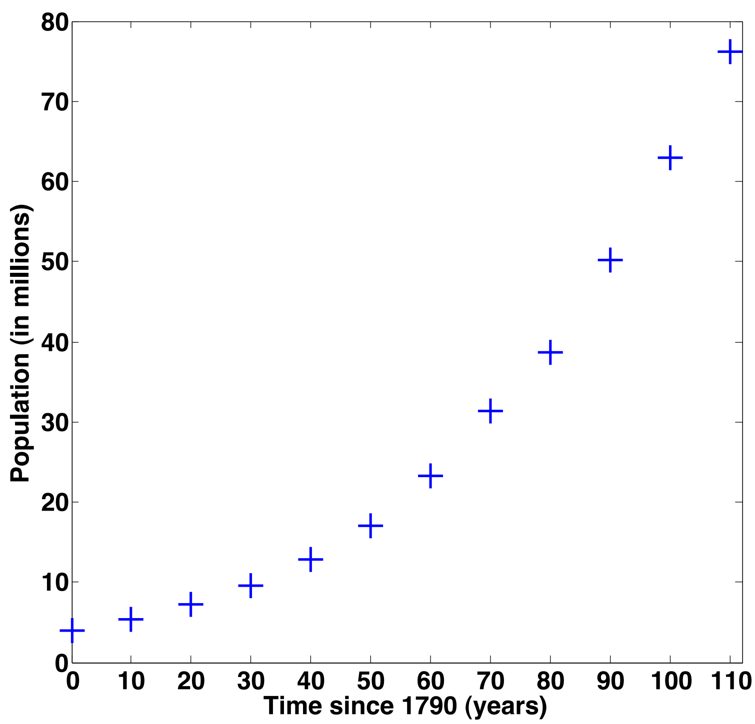
Suppose that you are presented with some data and want to figure out what type of function generated this data. For example, you may wish to better understand how the population of a certain region depends upon time, or you may wish to know how precise a particular method is for the numerical computation of derivatives.

In this project, we will learn to differentiate between data arising from a power function $f(t) = ct^k$ or an exponential function $g(t) = ce^{kt}$. Further, we will learn how to approximate the constants c and k from the data.

2. Case Study: U.S. Census Data

The table below provides population data (in millions) from 1790 - 1900 for the United States taken from http://www.census.gov/. It is not clear from the plot whether the data exhibits polynomial or exponential growth, or neither.

Year	Population (millions)
1790	3.9
1800	5.3
1810	7.2
1820	9.6
1830	12.9
1840	17.1
1850	23.2
1860	31.4
1870	38.6
1880	50.2
1890	63.0
1900	76.2

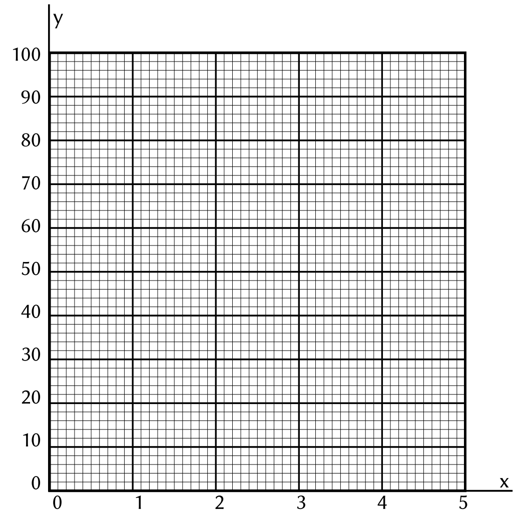


We begin by generating data from a known exponential function so that we can learn how to "discover" the original form of the function. Next, we will apply these techniques to real data.

3. Exponential functions

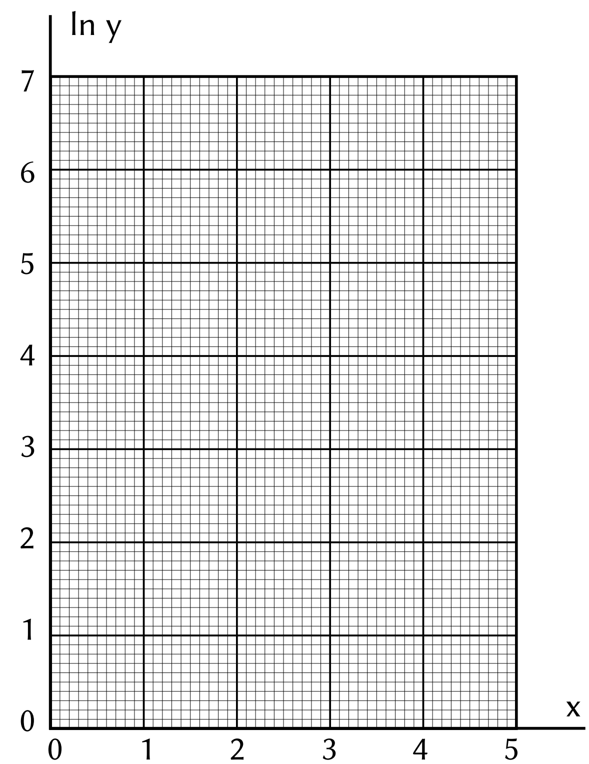
Consider the function $y = 3 \cdot 2^x$. We will use this function to generate some data. Fill in the table below for the indicated values of x. Next, plot the points (x, y) on the provided graph.

x	y
0	
1	
2	
3	
4	
5	



Now, using the values above and a calculator, fill in the table below and plot the points $(x, \ln y)$.

x	$\ln y$
0	
1	
2	
3	
4	
5	



4. What just happened?

You should notice that your second plot above appears to be a line. The second plot is usually called a "semilog plot" since it plots the natural logarithm of the y coordinate versus time. Should we have expected the semilog plot to give a line? The answer is yes since taking logs of both sides of the equation $y = 3 \cdot 2^x$ yields

$$\ln y = \ln 3 + \ln(2)x.$$

Thus, treating $(\ln y)$ as a variable, we see that there is a linear relationship.

5. More generally

In general, suppose that

$$y = c \cdot e^{kx}. \tag{12}$$

For example, if $y = 3 \cdot 2^x$, then we can write $y = 3 \cdot e^{(\ln(2))x}$. Taking ln of both sides of equation (12) yields

$$\ln(y) = \ln(c) + kx.$$

Thus we may conclude the following.

Fact: ***if*** $y = c \cdot e^{kx}$ ***for some*** c ***and*** k***, then the plot of*** x ***versus*** $ln(y)$ ***will be a line.***

More can even be concluded! Not only does the linearity of the semilog plot tell us that the function satisfies an exponential form, but it also gives us a method for determining c and k since

(1) $\ln(c)$ is the vertical intercept of the semilog plot when $x = 0$, and
(2) k is the slope of the semilog plot.

6. Power functions

Now suppose that we are dealing with a polynomial function such as

$$y = c \cdot x^k,$$

for some real numbers c and k. Will a similar idea work? Taking ln of both sides of the above equation yields

$$\ln(y) = \ln(c) + k\ln(x).$$

Therefore, we see that plotting $\ln(x)$ versus $\ln(y)$ will produce a line. Further, both c and k can be recovered by finding the slope and intercept of the resulting plot. Such a plot is referred to as a log-log plot since now both the dependent and independent variables have logarithms applied to them.

7. Problem

1. Consider the function

$$f(x) = e^{-x^2/2},$$

which plays an important role in probability. We will numerically approximate $f'(0.5)$ using three different methods:

(1) The forward difference: $f'(0.5) \approx \dfrac{f(0.5+h) - f(0.5)}{h}$.
(2) The backward difference: $f'(0.5) \approx \dfrac{f(0.5) - f(0.5-h)}{h}$.
(3) The centered difference: $f'(0.5) \approx \dfrac{f(0.5+h) - f(0.5-h)}{2h}$.

Fill in the following tables with the relevant data:

h	$\dfrac{f(0.5+h)-f(0.5)}{h}$	$\dfrac{f(0.5)-f(0.5-h)}{h}$	$\dfrac{f(0.5+h)-f(0.5-h)}{2h}$
1/5	-0.498 961		
1/10			
1/50			
1/100			
1/200			

The exact value of the derivative is $-0.44124845\cdots$. In the table below, give the error induced by the different methods

$$E(h) = \left|f'(0.5) - \text{approximate derivative with parameter } h\right|.$$

Note the absolute values above!

h	$E(h)$	$E(h)$	$E(h)$
1/5	0.057 713	0.073 746	0.008 016
1/10			
1/50			
1/100			
1/200			

Produce log-log plots of the data in the three tables above and derive the functional dependence of the error for each method as a function of h. Can you conclude that one method is definitively "better" then the others? In what precise sense is it better?

8. Report Instructions

Your project writeup should include:

(1) Summary of ideas detailed in project above including a derivation showing why the semilog or log-log plots will give a linear relationship under different circumstances.
(2) Complete solution to Exercise 7.1.
(3) Complete the following two exercises:
   (a) Taking the data from the US Census on the first page of this project, determine if the population data exhibits exponential or polynomial growth. Drawing a line by hand through the resulting linear plot (this will be either a log-log or semilog plot), determine the relevant constants c and k. That is, find a time dependent function of the population of the United States that fits the data well. What does the model predict the population of the United States would be in the year 2010? Explain.
   (b) A numerical method used to calculate integrals produced the following errors, $E(h)$, where $E(h)$ is a function of a certain parameter h:

h	$1/10$	$1/20$	$1/25$	$1/40$	$1/50$
$E(h)$	0.071	0.0174	0.0114	0.00438	0.0027

   A scientist believes that $E(h)$ is well approximated by a polynomial of the form Ch^k, for some $C, k > 0$, where k is a positive integer. Is the scientist right? Explain. If the scientist is right, what are the values of k and C?

Riemann Sums

1. Purpose of this project

We will attempt to gain an understanding of the definite integral

$$\int_a^b f(x)dx,$$

by computing different types of Riemann sums. We will learn that there is more than one such sum, but that one of them is much more accurate than the others.

2. Approximations

Consider the definite integral

$$\int_0^4 \sqrt{x+1}\, dx.$$

We will use a variety of Riemann sums to approximate this integral. First, we discretize $[0, 4]$ into N equally spaced intervals, $[x_{k-1}, x_k]$, with

$$\Delta x = x_k - x_{k-1} = \frac{4}{N}.$$

For example, in the case $N = 3$, we have $\{x_0, x_1, x_2, x_3\} = \{0, \frac{4}{3}, \frac{8}{3}, 4\}$ with associated intervals

$$[x_0, x_1] = [0, \tfrac{4}{3}], \quad [x_1, x_2] = [\tfrac{4}{3}, \tfrac{8}{3}], \quad [x_2, x_3] = [\tfrac{8}{3}, 4].$$

Letting

$$f(x) = \sqrt{x+1},$$

we will now compute sums of the form

$$\sum_{k=1}^{N} f(c_k)\Delta x,$$

for different values of N, and where $c_k \in [x_{k-1}, x_k]$ are the *sample points.*

See the lecture notes Chapter VII, Sections 1 and 2; also see the website http://mathworld.wolfram.com/RiemannSum.html for interactive pictures of Riemann sums.

3. Left sum

For the "Left sum", we will choose $c_k = x_{k-1}$, i.e. the left endpoint of the interval $[x_{k-1}, x_k]$, for all $k = 1, \dots, N$. For example, when $N = 3$, the left sum is

$$f(0)\tfrac{4}{3} + f(\tfrac{4}{3})\tfrac{4}{3} + f(\tfrac{8}{3})\tfrac{4}{3} = 1 \cdot \tfrac{4}{3} + \sqrt{\tfrac{7}{3}}\,\tfrac{4}{3} + \sqrt{\tfrac{11}{3}}\,\tfrac{4}{3} \approx 5.9231,$$

and when $N = 6$ the sum is

$$\sum_{k=1}^{6} f\left(\frac{4(k-1)}{6}\right)\frac{4}{6} \approx 6.3647. \tag{13}$$

Using your calculator, software such as Excel, or the Riemann-sum website http://mathworld.wolfram.com/RiemannSum.html, fill in the table below with the associated left sums.

N	$\sum_{k=1}^{N} f(x_{k-1})\Delta x$
6	6.3647
8	
12	
15	

4. Right sum

For the "Right sum", we will choose $c_k = x_k$, i.e. the right endpoint of the interval $[x_{k-1}, x_k]$, for all $k = 1, \dots, N$. For example, when $N = 3$, the right sum is

$$f(\tfrac{4}{3})\tfrac{4}{3} + f(\tfrac{8}{3})\tfrac{4}{3} + f(\tfrac{12}{3})\tfrac{4}{3} = \sqrt{\tfrac{7}{3}}\tfrac{4}{3} + \sqrt{\tfrac{11}{3}}\tfrac{4}{3} + \sqrt{\tfrac{15}{3}}\tfrac{4}{3} \approx 7.57126. \tag{14}$$

Using your calculator, or software, such as Excel, fill in the table below with the associated right sums.

N	$\sum_{k=1}^{N} f(x_k)\Delta x$
6	7.18877
8	
12	
15	

5. Midpoint sum

The actual value of the integral is

$$\int_0^4 \sqrt{x+1}\, dx = \frac{2}{3}\left(5\sqrt{5} - 1\right) \approx 6.786893. \tag{15}$$

In the previous sections of this project you should have found that the left sums always underestimated the correct solution for this problem while the right sum always overestimated. This should lead you to believe that perhaps it would be best to choose c_k to be the midpoint of x_{k-1} and x_k. That is, we suspect that a good approximation method would be

$$\sum_{k=1}^{N} f\left(\frac{x_{k-1} + x_k}{2}\right) \Delta x.$$

For example, when $N = 3$, the sum above sum is

$$f(\tfrac{2}{3})\ \tfrac{4}{3} + f(\tfrac{6}{3})\ \tfrac{4}{3} + f(\tfrac{10}{3})\ \tfrac{4}{3} \approx 6.80628.$$

Using your calculator, or Excel, fill in the table below with the associated midpoint sums.

N	$\sum_{k=1}^{N} f\left(\frac{x_{k-1}+x_k}{2}\right)\Delta x$
6	6.79193
8	
12	
15	

6. Report instructions

Your project should include

(1) Summary of ideas in project including all computations performed and conclusions reached. *Why did the left hand sum underestimate the true solution for this problem, while the right hand sum overestimated?* (Hint: draw some pictures!)

(2) Find formulas for x_k and c_k in the Left-sum (13). What is $\frac{4(k-1)}{6}$ doing there? Write the Right-sum (14) with $N = 3$ using summation notation in the manner of Equation (13).

(3) Do the integral that leads to the actual value stated in Equation (15).

(4) A function that is central to the study of probability is

$$f(x) = e^{-x^2/2}.$$

Compute the left hand, right hand, and midpoint sums for this function integrated over the interval $[0, 2]$ for several increasing values of N. *Do the left hand sums underestimate the integral, or do they overestimate the integral?* Stop each computation when you are confident that your N is large enough so that your value is accurate to two decimal places. How do you know you can stop when you did?

N	left sum	right sum	midpoint sum
6			
8			
12			
15			
...			
...			