# Unit V
Integration

These materials were developed for use at Grinnell College and neither Grinnell College nor the author, Mark Schneider, assume any responsibility for their suitability or completeness for use elsewhere nor liability for any damages or injury that may result from their use elsewhere.

Unit V
Integration

This activity guide works through some basic methods of numerical integration, and then takes a quick look at symbolic integration. Just as we saw in interpolation, and then again in the activity guide on finding roots, the tools of Taylor's series and passing polynomials through points are extremely useful. In this way, we can take functions that are difficult to deal with and reduce them to piecewise functions that are easy to deal with. We will see a few of the simplest forms of these, and then have a brief brush with danger as we take a glimpse of integration of functions that have nasty asymptotes, which are not well approximated by polynomials.

Guidebook Entry V.1: Riemann to Trapezoid in One Easy Step

We are introduced to the integral in calculus as the area under the curve, which is approximated as a Riemann sum. In the figure below, we show one view of such a sum to approximate the integral of the function from 1 to 7.
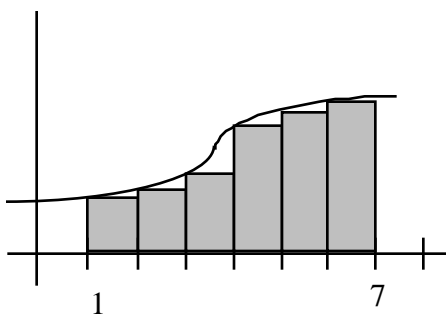
Does this sum over-estimate or under-estimate the value of the integral?

How would this change if the function were monotonically decreasing?

The sum uses the value of the function at the beginning of each sub-interval. How would this approximation change if you used the value of the function at the end of each sub-interval?

Can you suggest a better place to evaluate the function over the sub-interval to get a more accurate answer?

Now, let's make the next best approximation to the area under the curve. Rather than approximating each sub-interval as a rectangle, approximate each sub-interval as a trapezoid. This is equivalent to approximating the function over the sub-interval as a straight line from the function value at the beginning to the function value at the end. Write out the value of the area of the trapezoid of a single sub-interval.

Now write an expression for the full integral from 1 to 7 in the trapezoid approximation, visiting **function values f(x) only at integer x values**.

Let's get an idea of how well these approximations work with a function that we can integrate easily analytically to check our accuracy. Let's look at f(x) = 1+sin(x). What is the analytical value of the integral from 1 to 7?

Use Excel to calculate the approximate integral for each of the following cases:

Riemann, f at beginning of interval:

Riemann, f at end of interval:

Midpoint summation:

Trapezoid approximation:

How do each of these compare with the analytical value?

In the next approximation, known as Simpson's rule, we take the next step from the linear approximation of the trapezoid rule, and give each subinterval a midpoint value as well as the two endpoints, and pass a quadratic function through those three points. We then take the integral of the function underneath that curve. It sounds rather complicated, but you will find that in fact it is extremely easy to apply, once you have derived it. And as you might expect, since it is in fact one higher term in the Taylor's series (matching a second derivative instead of just a first), it will converge more quickly than the trapezoid, to one higher power of the sub-interval size.

Guidebook Entry V.2: Simpson's Rule

Before we can derive the Simpson's Rule integration rule, we need to find the quadratic curve that passes through three points. Rather than make them an arbitrary three points, let's place the first point at the origin, and the following two values, y_a and y_b at equally spaced x values, which we will call $x_a = $ and $x_b = 2$. Given this, calculate the coefficients of the quadratic equation $y = ax^2 + bx + c$ that passes through these three points $(0,0)$, $(,y_a)$, and $(2 ,y_b)$.

What is the value of the integral for an arbitrary function quadratic function $y = ax^2 + bx + c$ evaluated from zero to 2 ?

Now, substitute into your formula for the integral the values of a, b, and c as calculated above. Simplify your formula as best you can.

Does the absolute x value appear in this formula anywhere? Should it?

Now, let's gain some generality by moving these three points up by some amount y_o. We will thereby map the origin to $(0, y_o)$, the point (Δ, y_a) to $(\Delta, y_a + y_o)$, and $(2\Delta, y_b)$ to $(2\Delta, y_b + y_o)$. How much will this add to the integral?

Rewrite the expression for the integral in terms of these new y values y_o, $y_1 = y_o + y_a$ and $y_2 = y_o + y_b$.

Now imagine that we break our integral into a bunch of sub-intervals, each having a the two endpoints and a midpoint. Write a general algorithm for calculating the integral. What restriction is there on the number of points? Are the endpoints treated differently? Check your result with your instructor.

Now apply Simpson's rule to the same integral you performed above. Do you get a better value?

Higher order integration formulas abound, although for most purposes, the Simpson's rule will be adequate. The shortcomings of numerical integration routines are not in accuracy, in general, but in dealing with problematic functions. Really fancy algorithms don't do much more than fitting polynomials to sections of the function, but rather spend their attention on deciding how to space the sampled points. As you might imagine, you would want to have points spaced closely if the function has a lot of wiggles in it, and less closely if the funtion is flat and boring. Fancy algorithms allow the point spacing to be different over the range, and have ways of optimizing the choice of spacing. Virtually any reasonable book on numerical methods will talk about such techniques.

Guidebook Entry V.3: Using Mathematica for Integration

Mathematica will, you might have guessed, automatically perform a numerical integration. Try this on our same function. The form of the command to perform this integration is as follows:

NIntegrate[1+sin[x] , {x , 1 , 7 }]

Try this, and check your value against the your previous results.

Mathematica will also perform the integration symbolically. This is done through a command in the following format:

Integrate[1+sin[x] , {x , 1 , 7 }]

where this produces an analytic version of the definite integral. Try this, and compare the result to the indefinite integral:

Integrate[1+sin[x] , x] .

One can also get a numerical value for the analytical version of the definite integral as follows:

$$N[Integrate[1+sin[x] , \{x , 1 , 7 \}]] .$$

Describe the difference between this result and the result of using the NIntegrate command. If you're not sure, check with your instructor.

There are sometimes instances in which the polynomial approximations do not work well. For example, if you wish to make an integration that extends to infinity, it is difficult to choose the correct sub-interval size. Or if your function has a vertical or horizontal asymptote, it becomes problematic, since polynomials never have asymptotes. In the next example, which I have borrowed from Acton's book Numerical Methods that Work, we get a glimpse of how one can wrestle with nasty functions. You may do this final exercise for extra credit if you have the time.

Guidebook Entry V.4: Tricky Integration (Optional--for extra credit)

Let's imagine that we want to integrate the function

$$f(x) = \frac{1}{e^x - 1}$$

from nearly zero to one. First, graph this function to see that indeed it misbehaves at the origin. Sketch it below, or attach a graph.

This function can be integrated analytically (you might want to try Mathematica on it) to

$$\frac{dx}{e^x - 1} = \ln(e^x - 1) - x$$

which misbehaves if the integrand goes from zero. However, lets consider integrating from some point close to zero, say 0.01, as Acton chooses. Evaluate this integral from 0.01 to 1.

Now try using either trapezoid or Simpson on this integral, and see how close you can get with a reasonable number of points evaluated (say 20 or so).

We can improve this greatly, however, if we pull out a simple analytical part. Use the fact that e^x is approximately $1+x$ for x values very near zero to get the limiting behavior of this function near zero:

With this in mind, show that you can rewrite the integral as

$$\frac{1}{e^x - 1} - \frac{1}{x} \; dx + \; \frac{1}{x} dx \; .$$

Now graph the first function, and show that it no longer misbehaves at the origin.

The second function of course integrates easily to the natural logarithm, leaving our "unknown" integral a nicer one to integrate. Try this, using your trapezoid or Simpson routine on the same point spacing and adding the analytically obtained logarithm term from the second integral. Do you get a value that is much closer to the correct one?

Mathematica, you should know, is remarkably clever at dealing with these various infinities, although not bullet-proof. It also allows you to integrate out to infinity; you can simple use the word Infinity in place of a number for one of the limits of integration. You might want to try this if you have time.

1. In Unit V, you compared the performance of a beginning-of-subinterval Riemann sum, a midpoint sum, and a trapezoid sum. By what factor do each of these improve if you cut the sub-interval size in half?