# Numerical Integration

©Alan H. Stein
The University of Connecticut at Waterbury
stein@math.uconn.edu     http://www.math.uconn.edu/~stein

## Introduction

One frequently runs across definite integrals which cannot be calculated using the Fundamental Theorem of Calculus. In such cases, one resorts to ***numerical integration*** techniques in order to obtain an approximate value for the integral.

The crudest form of numerical integration is a Riemann Sum. In an elementary calculus course, we also learn about slightly more sophisticated techniques known as the Trapezoid Rule and Simpson's Rule.

For each method, we are interested in three questions:

(1) How do we carry out the numerical integration?
(2) How large is the error?
(3) How can we ensure that the error is smaller than a predetermined bound?

*As with all mathematics, this set of notes should be read with a pencil in one's hand and one should be drawing appropriate sketches as one reads.*

## Standard Notation

For each method, we will assume that we are trying to approximate an integral of the form $\int_a^b f(x)\,dx$, where $a < b$ and $f$ and as many of its derivatives as we need are continuous on the entire interval $[a, b]$.

We will also generally partition the interval $[a, b]$ into the partition $\mathcal{P} = \{x_0, x_1, x_2, \ldots, x_n\}$, where $h = \Delta x = \dfrac{b-a}{n}$ for some positive integer $n$ and $x_k = a + kh$ for $k = 0, 1, 2, \ldots, n$. We also let $y_k = f(x_k) \forall k$.

In deriving the various formulas, we will visualize $f$ as if it takes on only positive values and use the terms integral and area interchangeably. The formuas obtained are used even when the integral cannot be interpreted as an area.

## Riemann Sums

In general, a Riemann Sum is an expression of the form $\sum_{k=1}^n f(c_k)\Delta x_k$, where $c_k \in [x_{k-1}, x_k] \forall k$, without any specific requirement regarding the form of the partition used.

For the purposes of numerical integration, one would naturally use a partition where the intervals are all of equal size and choose the $c_k$ in an organized way.

If one lets $c_k$ be the left hand endpoint of the $k^{\text{th}}$ interval, one obtains the ***Left Hand Rule Approximation*** $L(f) = L(f, a, b) = \dfrac{b-a}{n} \sum_{k=1}^n y_{k-1}$.

If one lets $c_k$ be the right hand endpoint of the $k^{\text{th}}$ interval, one obtains the ***Right Hand Rule Approximation*** $R(f) = R(f, a, b) = \dfrac{b-a}{n} \sum_{k=1}^n y_k$.

If one lets $c_k$ be the midpoint of the $k^{\text{th}}$ interval, one obtains the ***Midpoint Rule Approximation*** $M(f) = M(f,a,b) = \dfrac{b-a}{n}\sum_{k=1}^{n} f\left(\frac{x_{k-1}+x_k}{2}\right)$.

**Example.** Consider $\int_1^2 \dfrac{1}{x}\,dx$. The value of the integral is, of course, $\ln 2 \approx 0.69314718056$. Suppose we use each of these three rules with $n = 100$.

$$L(1/x,1,2) = \frac{2-1}{100}\left(\frac{1}{1} + \frac{1}{1.01} + \frac{1}{1.02} + \frac{1}{1.03} + \cdots + \frac{1}{1.99}\right) = 0.69565343048.$$

$$R(1/x,1,2) = \frac{2-1}{100}\left(\frac{1}{1.01} + \frac{1}{1.02} + \frac{1}{1.03} + \frac{1}{1.04} + \cdots + \frac{1}{2}\right) = 0.69065343048.$$

$$M(1/x,1,2) = \frac{2-1}{100}\left(\frac{1}{\frac{1+1.01}{2}} + \frac{1}{\frac{1.01+1.02}{2}} + \frac{1}{\frac{1.02+1.03}{2}} + \cdots + \frac{1}{\frac{1.99+2}{2}}\right) = 0.69314405563.$$

*If one looks at the graph of $y = 1/x$, one can observe that the estimate obtained by the Left Hand Rule is an overestimate and the estimate obtained with the Right Hand Rule is an underestimate.*

## The Trapezoid Rule

One can view the various Riemann Sum approximations as being obtained by approximating portions of the graph of $f$ by horizontal lines. In the case of the Left Hand Rule, the portion of the graph for $x_{k-1} \le x \le x$ is approximated by the horizontal line through $(x_{k-1}, y_{k-1})$. In the case of the Right Hand Rule, the horizontal line through $(x_k, y_k)$ is used and in the case of the Midpoint Rule, the horizontal line through $\left(\dfrac{x_{k-1}+y_{k-1}}{2}, f\left(\frac{x_{k-1}+y_{k-1}}{2}\right)\right)$ is used.

Another natural line to use is the line joining $(x_{k-1}, y_{k-1})$ with $(x_k, y_k)$. That gives us a trapezoid, with bases $y_{k-1}$ and $y_k$, height $h = x_k - x_{k-1} = \dfrac{b-a}{n}$ and area $\dfrac{y_{k-1}+y_k}{2}\cdot\dfrac{b-a}{n}$.

If one adds all the areas together, one obtains $\sum_{k=1}^{n}\dfrac{y_{k-1}+y_k}{2}\cdot\dfrac{b-a}{n} = \dfrac{b-a}{2n}[(y_0+y_1) + (y_1+y_2)+(y_2+y_3)+\cdots+(y_{n-1}+y_n)] = \dfrac{b-a}{2n}(y_0+2y_1+2y_2+2y_3+\cdots+2y_{n-1}+y_n)$. This is the Trapezoid Rule. We will write $T(f) = T(f,a,b) = \dfrac{b-a}{2n}(y_0+y_1+y_2+y_3+\cdots+y_{n-1}+y_n)$.

**Example.** Let's again estimate $\int_1^2 \dfrac{1}{x}\,dx$ with $n = 100$.

We get $T(1/x,1,2) = \dfrac{2-1}{2\cdot 100}\left(\dfrac{1}{1} + 2\cdot\dfrac{1}{1.01} + \cdots + 2\cdot\dfrac{1}{1.99} + \dfrac{1}{2}\right) = 0.69315343048.$

## Simpson's Rule

It would intuitively seem reasonable that the area under a curve is better approximated by the area under a curve than by the area under a straight line. In a sense, the simplest possible curve is a parabola. We may think of a parabola as the graph of a second degree polynomial, with a straight line being the graph of a first degree polynomial.

Just as two points determine a line, three points determine a parabola. It should therefore be possible to determine the area of a parabola if one knows three points on the parabola.

Consider a parabola with equation $y = ax^2 + bx + c$ containing the points $(-h, y_{-h})$, $(0, y_0)$ and $(h, y_h)$. We can easily calculate $\int_{-h}^{h} ax^2 + bx + c\,dx = ax^3/3 + bx^2/2 + cx\,|_{-h}^{h} = 2ah^3/3 + 2ch = (2ah^2 + 6c)h/3$. We will eventually write this in terms of $h$, $y_{-h}$, $y_0$ and $y_h$.

Plugging $-h$, 0 and $h$ into the equation of the parabola, we obtain

$$y_{-h} = ah^2 - bh + c \tag{1}$$

$$y_0 = c \tag{2}$$

$$y_h = ah^2 + bh + c \tag{3}$$

Adding the first and third of these together, one obtains $y_{-h} + y_h = 2ah^2 + 2c$. We thus find $2ah^2 + 6c = 2ah^2 + 2c + 4c = y_{-h} + y_h + 4c = y_{-h} + y_h + 4y_0$ and can conclude $\int_{-h}^{h} ax^2 + bx + c\,dx = (y_{-h} + 4y_0 + y_h)h/3$.

Based on this, we approximate $\int_{x_0}^{x_2} f(x)\,dx$ by $(y_0+4y_1+y_2)\cdot\dfrac{b-a}{n}/3 = \dfrac{b-a}{3n}(y_0+4y_1+y_2)$.

Similarly, we approximate $\int_{x_2}^{x_4} f(x)\,dx$ by $\dfrac{b-a}{3n}(y_2 + 4y_3 + y_4)$.

Continuing, we can approximate the entire integral $\int_a^b f(x)\,dx$ by $\dfrac{b-a}{3n}(y_0 + 4y_1 + y_2) + \dfrac{b-a}{3n}(y_2+4y_3+y_4)+\cdots+\dfrac{b-a}{3n}(y_{n-2}+4y_{n-1}+y_n) = \dfrac{b-a}{3n}[(y_0+4y_1+y_2)+(y_2+4y_3+y_4)+\cdots+(y_{n-2}+4y_{n-1}+y_n)] = \dfrac{b-a}{3n}(y_0+4y_1+2y_2+4y_3+2y_4+4y_5+\cdots+2y_{n-2}+4y_{n-1}+y_n)$.

This is called Simpson's Rule and we write $S(f, a, b) = \dfrac{b-a}{3n}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + 4y_5 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n)$.

**Example.** Let's again estimate $\int_1^2 \dfrac{1}{x}\,dx$ with $n = 100$.

We get $S(1/x, 1, 2) = \dfrac{2-1}{3 \cdot 100}\left(\dfrac{1}{1} + 4\dfrac{1}{1.01} + 2\dfrac{1}{1.02} + \cdots + 4\dfrac{1}{1.99} + \dfrac{1}{2}\right) = 0.69314718087$.

## Error Estimates

Let $E_T$, $E_M$ and $E_S$ represent the errors in the Trapezoid, Midpoint and Simpson's Rules, respectively, and let $K$ be a bound on $|f''(x)|$ and $K^*$ a bound on $|f^{(4)}(x)|$ on $[a, b]$. We have the following bounds on those errors.

$$|E_T| \leq \frac{K(b-a)^3}{12n^2} \tag{4}$$

$$|E_M| \leq \frac{K(b-a)^3}{24n^2} \tag{5}$$

$$|E_S| \leq \frac{K^*(b-a)^5}{180n^4} \tag{6}$$

*We will shortly get an idea of how bounds such as these may be derived. It's interesting to note that the bound using the Midpoint Rule is actually about twice as good as that for the Trapezoid Rule. This corresponds to the approximations obtained above, where the error using the Midpoint Rule is approximately 0.00000312493, about half the size of the error of approximately 0.00000624992 using the Trapezoid Rule. The error obtained using Simpson's Rule, of approximately 0.00000000031, is far better than that obtained using either the Trapezoid Rule or the Midpoint Rule.*

For the integral being estimated, $f(x) = 1/x = x^{-1}$, so $f'(x) = -x^{-2}$, $f''(x) = -2x^{-3}$, $f'''(x) = 6x^{-4}$ and $f^{(4)}(x) = -24x^{-5}$. On the interval $[1, 2]$, $|f''(x)| = 2/x^3 \leq 2$, so we may take $K = 2$. Similarly, $|f^{(4)}(x)| = 24/x^5 \leq 24$, so we may take $K^* = 24$. We thus obtain

$$|E_T| \leq \frac{K(b-a)^3}{12n^2} \leq \frac{2(2-1)^3}{12 \cdot 100^2} \approx .000016666 \tag{7}$$

$$|E_M| \leq \frac{K(b-a)^3}{24n^2} \leq \frac{2(2-1)^3}{24 \cdot 100^2} \approx .000008333 \tag{8}$$

$$|E_S| \leq \frac{K^*(b-a)^5}{180n^4} \leq \frac{24(2-1)^5}{180 \cdot 100^4} \approx .000000001333. \tag{9}$$

Note that in each case, the actual error we obtained was less than the theoretical bound calculated here.

## Determining a Value for $n$

Suppose we want to ensure that the error in our approximation will be no greater than a predetermined value $\epsilon$. For each rule, we have a formula for a bound on the error involving the endpoints (which we know), a bound on a derivative (which we hopefully can calculate) and $n$. If we can ensure that the bound is less than $\epsilon$, it follows that the error will be less than $\epsilon$. In each case, ensuring the bound is less than $\epsilon$ amounts to solving an inequality for $n$.

Consider the integral $\int_1^2 1/x \, dx$ again. We have the bounds

$$|E_T| \leq \frac{K(b-a)^3}{12n^2} \leq \frac{2(2-1)^3}{12n^2} = \frac{1}{6n^2} \tag{10}$$

$$|E_M| \leq \frac{K(b-a)^3}{24n^2} \leq \frac{2(2-1)^3}{24 \cdot n^2} = \frac{1}{12n^2} \tag{11}$$

$$|E_S| \leq \frac{K^*(b-a)^5}{180n^4} \leq \frac{24(2-1)^5}{180 \cdot n^4} = \frac{2}{15n^4}. \tag{12}$$

Suppose we want to have an approximation correct to about 20 decimal places. We can do that by letting $\epsilon = 10^{-20}$. If we are using the Trapezoid Rule, we will need to ensure that $\dfrac{1}{6n^2} \leq 10^{-20}$. Solving for $n$, we get $n^2 \geq 10^{20}/6$, $n \geq \sqrt{10^{20}/6} \approx 4{,}082{,}482{,}904.64$. We can safely let $n = 4{,}082{,}482{,}906$, or approximately 4 billion.

From the Mean Value Theorem, one knows both that $f(x) - f(a) = f'(\xi_x)(x-a)$ and $m = \dfrac{f(b)-f(a)}{b-a} = f'(\xi)$ for some $\xi_x$ between $a$ and $x$ and some $\xi$ between $a$ and $b$. Thus, $E = \int_a^b (f'(\xi_x) - f'(\xi))(x-a)\,dx$.

The same Mean Value Theorem enables us to write $f'(\xi_x) - f'(\xi) = f''(\xi^*)(\xi_x - \xi)$ for some $\xi^*$ between $\xi$ and $\xi_x$. We can thus write $E = \int_a^b f''(\xi^*)(\xi_x - \xi)(x-a)\,dx$. If we know $|f''(x)| \le K$ on $[a,b]$, and recognizing that $|\xi_x - \xi| \le b-a$ since both $\xi$ and $\xi_x$ are between $a$ and $b$, it follows that $|E| \le \int_a^b |f''(\xi^*)(\xi_x - \xi)(x-a)|\,dx \le K(b-a)\int_a^b (x-a)\,dx = K(b-a)^3/2$.

Once again, we could do the same analysis on a subinterval $[x_{k-1}, x_k]$ and get essentially the same estimate, with the width $\dfrac{b-a}{n}$ in place of $b-a$. Thus, each subinterval can contribute no more than $K\left(\dfrac{b-a}{n}\right)^3/2 = \dfrac{K(b-a)^3}{2n^3}$.

Since there are $n$ such subintervals, the total possible error is $n$ times as great, or $\dfrac{K(b-a)^3}{2n^2}$.

*Remember, this is just a demonstration. Much better bounds can be obtained; the estimate in the text is one-sixth this size.*