# Numerical Integration

MAT 229, Spring 2021

Week 6

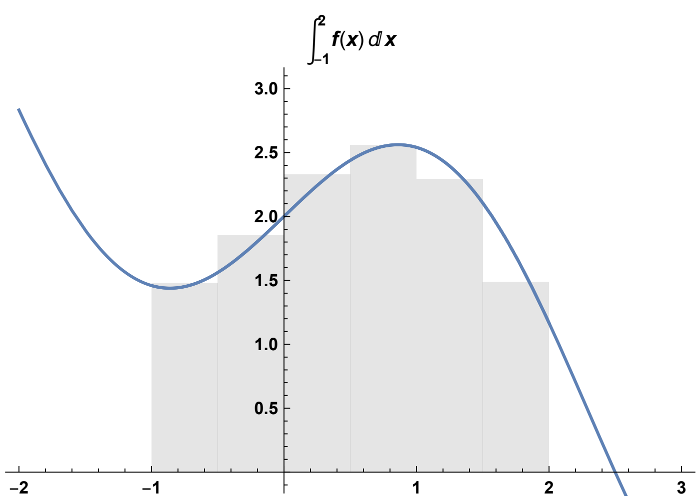
## Supporting materials

If you wish to get a different perspective on the notes below, try either of the following textbook sections.

- Stewart's *Calculus*
  Section 7.7: Approximate integration
- Boelkins/Austin/Schlicker's Active Calculus:
  Section 5.6: Numerical integration

## Rectangles

Consider the integral $\int_a^b f(x)\,dx$. To approximate the area under the curve $y = f(x)$ with rectangles of equal width, two things are needed:




- The number of rectangles.
- Where each rectangle should touch the curve, i.e. which value of $x$ to use in the function evaluation (rectangle height) in each subinterval.

## Rectangle base points

The entire width of this region is $b - a$. If we decide on $n$ rectangles, the width of each rectangle is

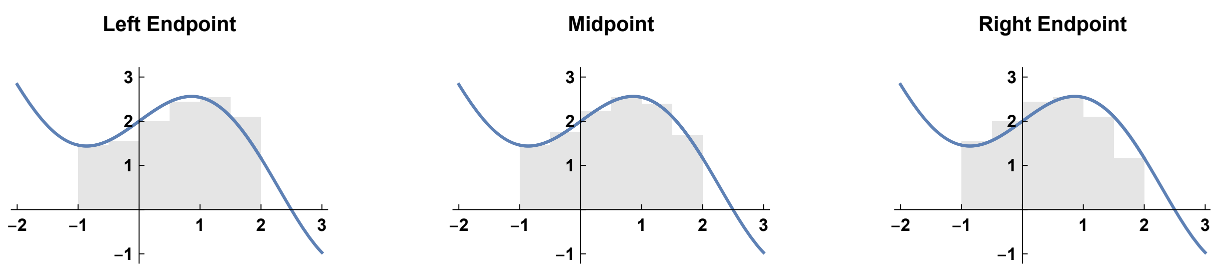
$$\Delta x = \frac{b-a}{n}$$

The locations of the base of each rectangle along the $x$-axis:

- The first rectangle on the left end starts at the lower limit of the integral and goes for one width: $x_0 = a$, $x_1 = a + \Delta x$
- The second rectangle begins where the first one ended and goes for another width: $x_1 = a + \Delta x$, $x_2 = a + 2\,\Delta x$
- The third rectangle begins where the second one ended and goes for another width: $x_2 = a + 2\,\Delta x$, $x_3 = a + 3\,\Delta x$
- Continue the process until you reach the $n^{\text{th}}$ rectangle: $x_{n-1} = a + (n-1)\,\Delta x$, $x_n = a + n\,\Delta x = b$ (Why is the last one $b$?)

## Rectangles: heights given by function values, chosen from the sub-intervals

Three standard ideas about where to sample the function lead to three different rectangle rules:  left endpoint, the midpoint, and the right endpoint.




So the approximate area of a method will be a sum of little rectangular areas that look like $f(x^*)\Delta x$, height*width, one for each of the $n$ subintervals. $x^*$ is a particular choice of $x$ within each subinterval (e.g. left endpoint, midpoint, right endpoint -- but other choices are possible, although we won’t get into them).

## Question

Do you see that the above three rules translate to the following formulas? One thing we typically do is “distribute out” a Δx, since the width is the same for each rectangle.

- Left endpoint rule: $\int_a^b f(x)\,dx \approx \Delta x\left(\sum_{k=1}^{n} f(a + (k-1)\,\Delta x)\right) = \Delta x\left(f(a) + f(a + \Delta x) + \ldots + f(a + (n-1)\,\Delta x)\right)$
- Midpoint rule: $\int_a^b f(x)\,dx \approx \Delta x\left(\sum_{k=1}^{n} f\left(\frac{x_{k-1}+x_k}{2}\right)\right) =$
$$\Delta x\left(\sum_{k=1}^{n} f\left(a + \left(k - \tfrac{1}{2}\right)\Delta x\right)\right) = \Delta x\left(f\left(a + \tfrac{1}{2}\,\Delta x\right) + f\left(a + \tfrac{3}{2}\,\Delta x\right) + \ldots + f\left(a + \left(k - \tfrac{1}{2}\right)\Delta x\right)\right)$$

- Right endpoint rule:
  $\int_a^b f(x)\,dx \approx \Delta x(\sum_{k=1}^{n} f(a + k\,\Delta x)) = \Delta x(f(a + \Delta x) + \ldots + f(a + (n-1)\,\Delta x) + f(a + n\,\Delta x))$

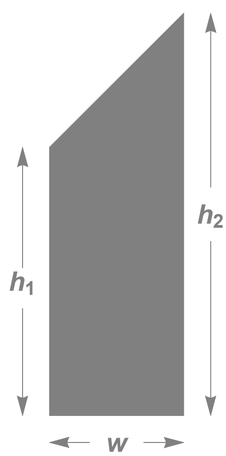
---

# Trapezoid rule

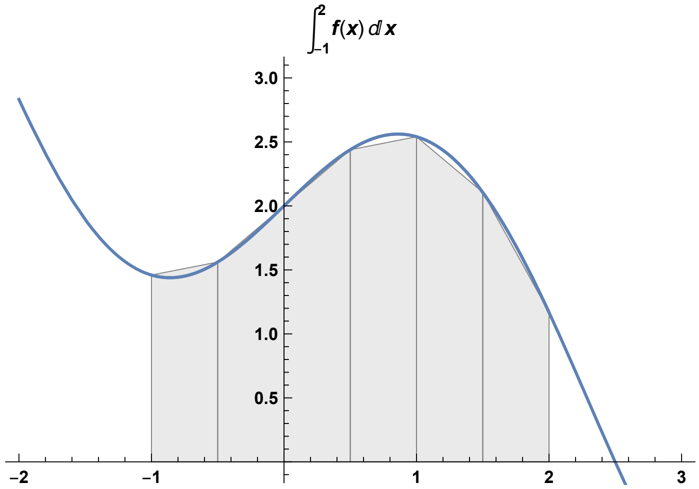
The trapezoid rule for estimating $\int_a^b f(x)\,dx$ with $n$ trapezoids is the average of the left endpoint rule and the right endpoint rule, each with $n$ rectangles. We can either
a. calculate Left and Right separately, and then average those two; or we can
b. average their **formulas**: Let $\Delta x = \frac{b-a}{n}$ and $x_k = a + k\,\Delta x$.

$$\int_a^b f(x)\,dx \approx \frac{1}{2}$$
$$((f(x_0) + f(x_1) + f(x_2) + \ldots + f(x_{n-2}) + f(x_{n-1}))\,\Delta x + (f(x_1) + f(x_2) + f(x_3) + \ldots + f(x_{n-1}) + f(x_n))\,\Delta x$$
$$= \frac{1}{2}\,\Delta x(f(x_0) + 2\,f(x_1) + \ldots + 2\,f(x_{n-1}) + f(x_n))$$
$$= \frac{1}{2}\,\Delta x\left(f(a) + 2\sum_{k=1}^{n-1} f(a + k\Delta x) + f(b)\right)$$

A trapezoid has the following shape and its area is $\frac{(h_1+h_2)\,w}{2}$




The term “trapezoid rule” is used for this rule because the above average of the left and right endpoint rules is equivalent to summing the areas of trapezoids -- trapezoids which I hope you recognize as the average of two rectangles!




## Example: you don't need a function if you have regularly spaced data.

A functional expression isn't necessary if you have regularly spaced data:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **t (hours)** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| **p (t) = proportion in *rigor mortis*** | $\frac{2}{123}$ | $\frac{16}{123}$ | $\frac{47}{123}$ | $\frac{61}{123}$ | $\frac{81}{123}$ | $\frac{92}{123}$ | $\frac{99}{123}$ | $\frac{103}{123}$ | $\frac{110}{123}$ | $\frac{111}{123}$ | $\frac{111}{123}$ | $\frac{113}{123}$ |

So while we may think "$p(7) = \frac{92}{123}$", there's really nothing to compute with. We just use the data. To estimate the average value of the underlying proportion $p(t)$, $2 \le t \le 13$, estimate the interval $\frac{1}{13-2} \int_2^{13} p(t)\,dt$ with the trapezoid rule on the data using $n = 11$.
And while it might be better to speak of t rather than x, and "Δt" rather than Δx, we'll continue to use x and Δx:

- $\Delta x = \frac{13-2}{11} = 1$
- $x_0 = 2, x_1 = 3, x_2 = 4, \ldots, x_{11} = 13$

$$\frac{1}{11} \int_2^{13} p(t)\,dt \approx \frac{1}{11} \left(\frac{1}{2} \Delta x (p(x_0) + 2\,p(x_1) + \ldots + 2\,p(x_{10}) + p(x_{11}))\right)$$

$$= \frac{1}{2}(1)\,\Big(\frac{2}{123} + 2\Big(\frac{16}{123} \;\Big|\; + \frac{47}{123} + \frac{61}{123} + \frac{81}{123} + \frac{92}{123} + \frac{99}{123} + \frac{103}{123} + \frac{110}{123} + \frac{111}{123} + \frac{111}{123}\Big) + \frac{113}{123}\Big)$$

$$= \frac{1}{22}\left(\frac{1}{123}\right)(2 + 32 + 94 + 122 + 162 + 184 + 198 + 206 + 220 + 222 + 222 + 113)$$

$$= \frac{1}{246}(1777) \approx 0.656689$$

# Technology

For values of $n$ other than small ones, using these numerical techniques by hand is tedious. However, for computing devices it is a piece of cake. Two things make the process easier to implement.

- A simple way to represent a function, as in the function notation $f(x) = x\,e^{-2x}$.

```
In[206]:= f[x_] := x E^(-2 x)
```

- A simple way to implement the summation notation: $\sum_{k=1}^{n} g(k) = g(1) + g(2) + g(3) + \ldots + g(n-1) + g(n)$.

## Example

Consider the integral $\int_2^6 \sqrt{x^3 - 5}\, dx$.

- Use the midpoint rule with $n = 50$ to approximate it. $\int_a^b f(x)\, dx \approx \Delta x \left(\sum_{k=1}^{n} f\left(a + \left(k - \frac{1}{2}\right) \Delta x\right)\right)$
- Estimate the error using the midpoint rule error estimate, absolute error $\leq \frac{K(b-a)^3}{24\, n^2}$ where $K \geq |f''(x)|$ for $a \leq x \leq b$.

## Question

Using the trapezoid rule, estimate $\int_1^4 x^{x/2}\, dx$ with error less than 0.001.

The error estimate for the trapezoid rule is absolute error $\leq \frac{K(b-a)^3}{12\, n^2}$ where $K \geq |f''(x)|$ for $a \leq x \leq b$.

Define the function and find a value for $K$.

(Work the problem, then compare your results with my results by clicking on the ⟫ button above.)

---

# Simpson’s rule

The error estimates for the trapezoid rule and the midpoint rule are similar.

- Trapezoid rule error estimate: absolute error $\leq \frac{K(b-a)^3}{12\, n^2}$
- Midpoint rule error estimate: absolute error $\leq \frac{K(b-a)^3}{24\, n^2}$

The midpoint rule looks to be about twice as accurate as the trapezoid rule using the same $n$. Simpson's rule for estimating $\int_a^b f(x)\, dx$ is the weighted average of the midpoint rule (counted twice) and the trapezoid rule (counted once), but computed with $m = 2\, n$ subintervals.

$$\text{Simpsons}_m = \frac{2\,\text{Midpoint}_n + \text{Trapezoid}_n}{3}$$

This is the easy way to compute Simpson’s rule! However, you could perform this weighted average on the rules themselves:

Let $\Delta x = \frac{b-a}{m}$ and $x_k = a + k\, \Delta x$.

$$\int_a^b f(x)\, dx \approx \frac{1}{3} \Delta x (f(x_0) + 4\, f(x_1) + 2\, f(x_2) + 4\, f(x_3) + 2\, f(x_4) + \ldots + 2\, f(x_{2n-2}) + 4\, f(x_{2n-1}) + f(x_{2n}))$$

Note that except for the first and last terms in this sum

- the odd indexed terms have a coefficient of 4: index = $2\,k - 1$
- the even indexed terms have a coefficient of 2: index = $2\,k$

In summation notation, Simpson's rule is $\int_a^b f(x)\,dx \approx \frac{1}{3}\Delta x\left(f(a) + 4\sum_{k=1}^{n} f(x_{2k-1}) + 2\sum_{k=1}^{n-1} f(x_{2k}) + f(b)\right)$

## Example

Consider the integral $\int_2^6 \sqrt{x^3 - 5}\,dx$.

- Use Simpson's rule with $m = 2n = 100$ (so that $n = 50$) to approximate it. $\int_a^b f(x)\,dx \approx \frac{1}{3}\Delta x\left(f(a) + 4\sum_{k=1}^{n} f(a + (2k-1)\Delta x) + 2\sum_{k=1}^{n-1} f(a + 2k\Delta x) + f(b)\right)$
- Estimate the error using $S_m$ , absolute error $\leq \frac{K(b-a)^5}{180\,m^4}$ where $K \geq \left|f^{(4)}(x)\right|$ for $a \leq x \leq b$.

## Question

Using Simpson's rule, estimate $\int_1^4 x^{x/2}\,dx$ with error less than 0.001.

The error estimate for Simpson's rule $S_m$ absolute error $\leq \frac{K(b-a)^5}{180\,m^4}$ where $K \geq \left|f^{(4)}(x)\right|$ for $a \leq x \leq b$. Define the function and find a value for $K$. Notice that we are thinking of *m=2n*, an even number. Find the corresponding value of m to assure an error less than 0.001.

(Work the problem, then compare your results with my results by clicking on the ⟩⟩ button above.)

---

# Perfect integration from Numerical Schemes

Because of the form of the error bounds for the Trapezoidal, Midpoint , and Simpson's rules, you should realize that the Trapezoidal and Midpoint rules get integrals of linear functions exactly right; and that Simpson's rule gets integrals of **cubic** polynomials exactly right. No matter how many subintervals you take!

Let's see.... The fewest subintervals Simpson's can use is 2. So take m=2 (then n=1 for the Trapezoidal and Midpoint rules).

## Example

Evaluate the definite integral $\int_0^3 (16 + x^2 - x^3)\,dx$

In[258]:=
```
a = 0.0
b = 3.0
n = 1
dx = (b - a) / n
f[x_] := 16 + x^2 - x^3
Plot[f[x], {x, 0, 3}]
```

In[264]:=
```
Integrate[f[x], {x, 0, 3}]
true = N[%]
```

Out[264]= $\frac{147}{4}$

```
In[266]:= lrrN = dx * Sum[f[a + (k - 1) dx], {k, 1, n}]
lrrN - true (* dramatic overestimate)
```

```
In[267]:= rrrN = Sum[dx * f[a + (k - 0) dx], {k, 1, n}]
rrrN - true (* dramatic underestimate)
```

```
In[270]:= midN = dx * Sum[f[a + (k - 1 / 2) dx], {k, 1, n}]
midN - true
```

```
In[268]:= trapN = 1 / 2 (lrrN + rrrN)
trapN - true
```

## Error estimate

For the trapezoid rule the absolute error is less than or equal to

$$\frac{K(b-a)^3}{12\, n^2}$$

```
In[∘]:= Plot[Abs[f''[x]], {x, a, b}]
```

```
In[∘]:= K2 = Abs[f''[b]]
```

```
In[∘]:= trapErr = K2 (b - a) ^3 / (12 * n^2)
```

```
In[∘]:= midErr = K2 (b - a) ^3 / (24 * n^2)
```

## Estimate and Error estimate for Simpson’s Rule:

```
In[272]:= S2N = (2 * midN + trapN) / 3
true
S2N - true
```

Perfect! From a bunch of crummy estimates, Simpson’s got it exactly right! How?

For Simpson’s rule the absolute error is less than or equal to

$$\frac{K(b-a)^5}{180\, n^4}$$

```
In[275]:= Plot[Abs[f''''[x]], {x, 0, 3}] (* identically zero! *)
```

```
In[276]:= K4 = 0
(* Since the fourth derivative is identically zero! *)
```

```
In[277]:= K4 (b - a) ^5 / (180 * (2 n) ^4)
(* = 0; Simpson's gets it exactly right! *)
```