

Lecture 7 : Approximate integration.

Given a definite integral  $\int_a^b f(x) dx$  there are 2 reasons we may not be able to solve it exactly:

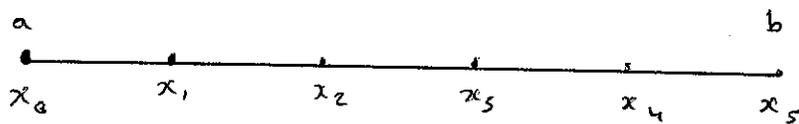
1) We cannot find an easy expression for an anti-derivative.

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

2) The function  $f(x)$  from a scientific experiment, hence there may be no easy formula for the function.

Today we will study various methods to approximate a definite integral.

Reminder: Recall that  $\int_a^b f(x) dx$  is defined as a limit of Riemann sums. Let's recall the definition of a Riemann sum (of length  $n$ ). First divide  $[a, b]$  into  $n$  subintervals of equal length  $\Delta x = \frac{b-a}{n}$ . Let  $x_i = a + i \Delta x$ , where  $0 \leq i \leq n$ .

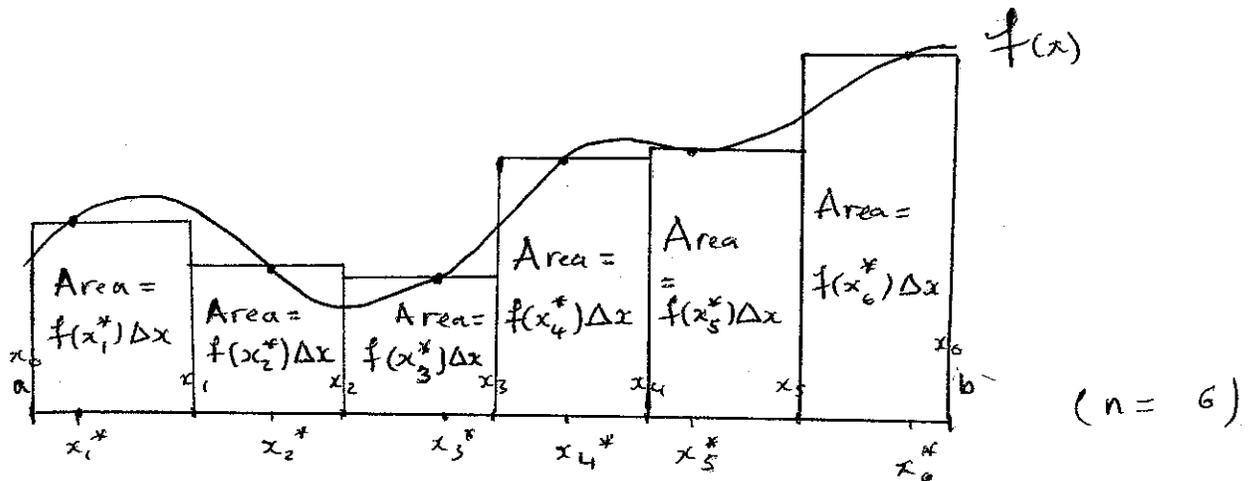


( $n = 5$  in picture)

Choose a point  $x_i^*$  in the  $i$ th subinterval  $[x_{i-1}, x_i]$ .

This data gives the Riemann sum:

$$\sum_{i=1}^n f(x_i^*) \Delta x.$$



$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) \Delta x$$

Hence a Riemann sum is an approximation to  $\int_a^b f(x) dx$ .

Easiest Approximations :

Left Endpoint Approximation : In Riemann sum choose

$$x_i^* = x_{i-1}, \text{ and write } L_n = \sum_{i=1}^n f(x_{i-1}) \Delta x$$

Right Endpoint Approximation : In Riemann sum choose

$$x_i^* = x_i, \text{ and write } R_n = \sum_{i=1}^n f(x_i) \Delta x$$

Midpoint Approximation : In Riemann sum choose

$$x_i^* = \bar{x}_i = \frac{1}{2} (x_{i-1} + x_i) = \text{midpoint of } [x_{i-1}, x_i] \text{ and}$$

$$\text{write } M_n = \sum_{i=1}^n f(\bar{x}_i) \Delta x$$

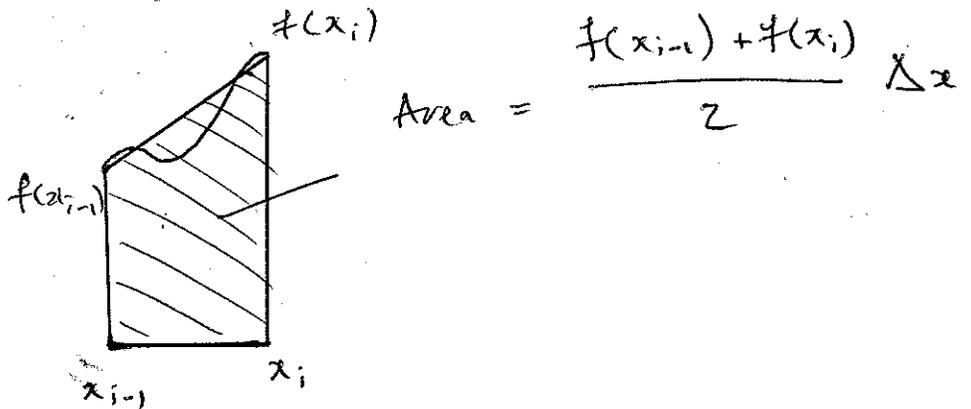
More interesting approximation:

Trapezoid Approximation Take the average of the

left and right endpoint approximation, i.e.

$$T_n = \frac{1}{2} (L_n + R_n) = \frac{1}{2} \left( \sum_{i=1}^n f(x_{i-1}) \Delta x + \sum_{i=1}^n f(x_i) \Delta x \right)$$
$$= \sum_{i=1}^n \left( \frac{f(x_{i-1}) + f(x_i)}{2} \right) \Delta x = \frac{\Delta x}{2} (f(x_0) + 2f(x_1) + 2f(x_2) + \dots$$
$$\dots + 2f(x_{n-1}) + f(x_n))$$

It's called the Trapezoid approximation (or rule) because of the following picture:

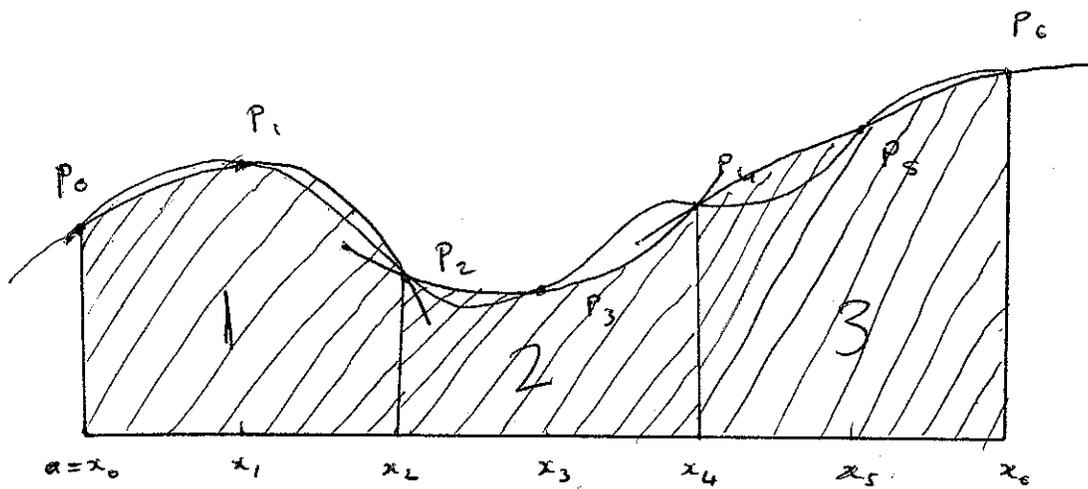


So in the sum we are adding areas of trapezoids and not rectangles.

Even more interesting approximation:

Instead of using straight line segments, use parabolas to approximate the curve. A parabola is just a polynomial of degree 2 or less. Here is a picture to illustrate the

idea:



Shaded area 1 = area under the parabola passing through  $P_0, P_1, P_2$ , between  $x_0$  and  $x_2$

Shaded area 2 = area under the parabola passing through  $P_2, P_3, P_4$ , between  $x_2$  and  $x_4$

Shaded area 3 = area under the parabola passing through  $P_4, P_5, P_6$ , between  $x_4$  and  $x_6$

Approximation is Area 1 + Area 2 + Area 3.

This is a good plan because it should be more accurate than trapezoid rule and we can work out each shaded area using integration of quadratic functions

### Simpson's Approximation (n Rule)

For  $n$  even,  $n/2$

$$S_n = \sum_{i=1}^{n/2} \left( \text{Area under parabola passing through } P_{2i-2}, P_{2i-1}, P_{2i}, \text{ between } x_{2i-2} \text{ and } x_{2i} \right)$$

$$= \frac{\Delta x}{3} (f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + \dots$$

↑

$$\dots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)$$

Some clever

algebra (see p 512)

For each approximation we have an error:

$$\int_a^b f(x) dx - L_n = E_L, \quad \int_a^b f(x) dx - R_n = E_R,$$

$$\int_a^b f(x) dx - M_n = E_M, \quad \int_a^b f(x) dx - T_n = E_T,$$

$$\int_a^b f(x) dx - S_n = E_S.$$

For larger  $n$ , these errors will get smaller and smaller. We want to find a balance between the size of the error and the size of  $n$  (ie the number of computations).

The Midpoint, Trapezoid and Simpson approximations have the following bounds on their errors:

Midpoint Error Bound      Suppose  $|f''(x)| \leq K$  for  $a \leq x \leq b$ .

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

Trapezoid Error Bound Suppose  $|f''(x)| \leq K$  on

$$a \leq x \leq b, \text{ then } |E_T| \leq \frac{K(b-a)^3}{12n^2}$$

Note that the error bound for the midpoint approximation is half that of the trapezoid approximation. Hence midpoint approximation is generally more accurate than trapezoid approximation.

Simpson Error Bound Suppose  $|f^{(4)}(x)| \leq K$  for

$$a \leq x \leq b, \text{ Then } |E_S| \leq \frac{K(b-a)^5}{180n^4}$$

Because there is an  $n^4$  in the denominator this bound is generally smaller than the ones for midpoint/trapezoid approximations. Hence Simpson's approximation is generally the most accurate. The downside is that we'd need to work out a 4<sup>th</sup> derivative.

For good examples of these approximations being used, look at Examples 2 (page 510), 3 (page 511) and 6 (page 515) of the text book.