Area under a curve

Objective:

To estimate the approximate area under the graph of a continuous function f(x) on the interval [a,b].

Basic approach:

Cover, or tessellate, the region with "tiles" of known area. Total area of tiles gives the required approximation.

To find area under curves, we use rectangular tiles.

Strategy:

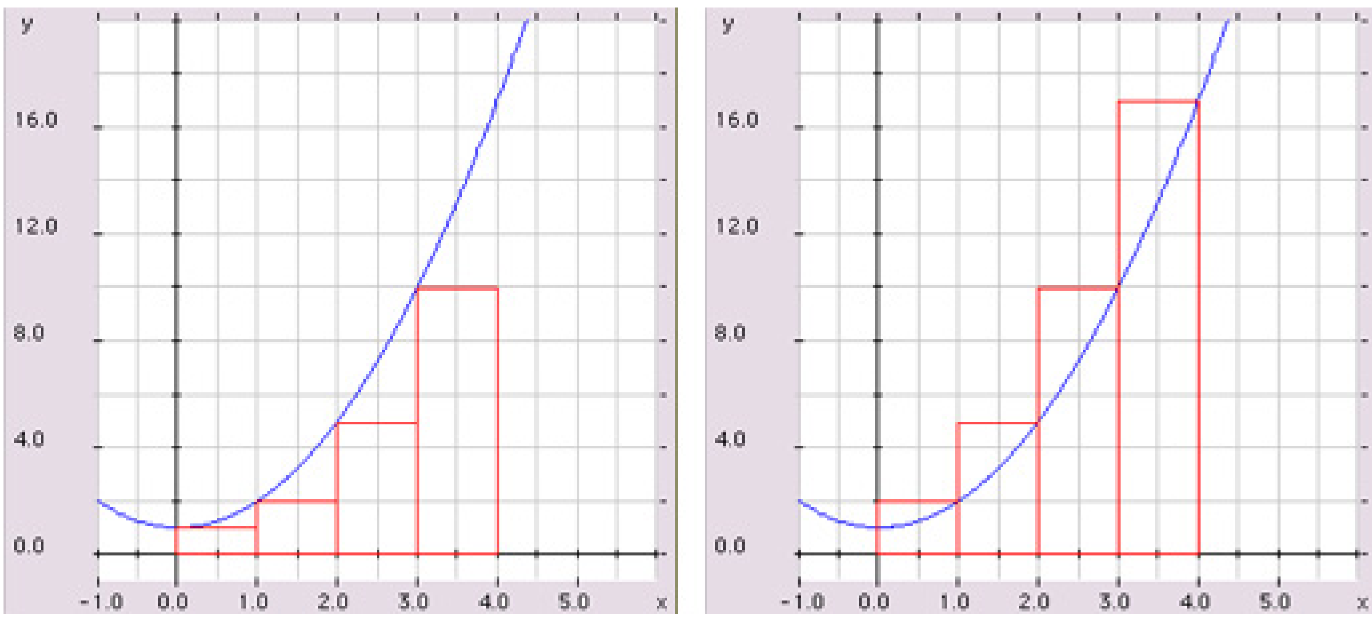
[1] Divide the given interval [a,b] into smaller pieces (sub-intervals).

[2] Construct a rectangle on each sub-interval & "tile" the whole area.

[3] Calculate total area of all the rectangles to get approximate area under f(x).

Example

Approximate the area under $y = f(x) = x^2 + 1$, from $x=0$ to 4, using 4 rectangles.



Solution:

Four intervals of equal width $\Rightarrow \Delta x = (4.0 - 0)/4 = 1.0$

Find x-values at interval boundaries: $x_0=0$, $x_1=1.0$, $x_2=2.0$, $x_3=3.0$, $x_4=4.0$.

To get rectangle heights, find $f(x)$ values at interval boundaries: $f(x_k) = x_k^2 + 1$.

$f(x_0)=1.0$, $f(x_1)=2.0$, $f(x_2)=5.0$, $f(x_3)=10.0$, $f(x_4)=17.0$.

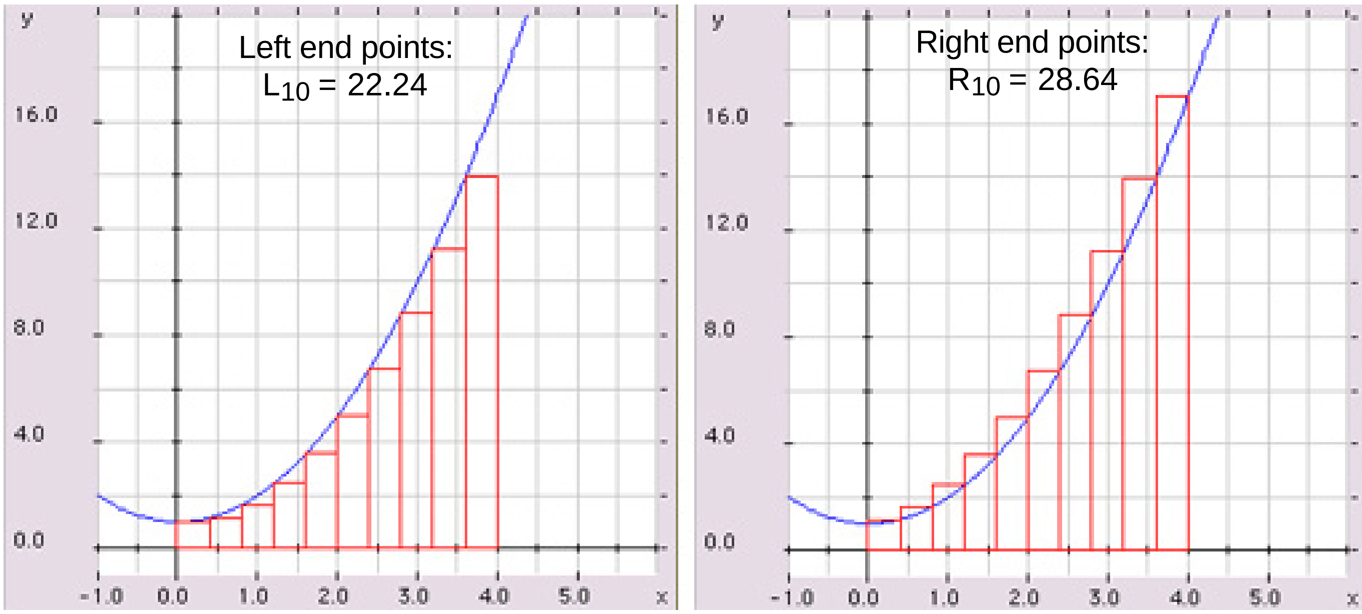
If we use the left end-points: Area = $\left[\sum_{k=0}^{3} f(x_k)\right]\Delta x$ = (1+2+5+10)*1.0 =18.0

If we use right end-points: Area = $\left[\sum_{k=1}^{4} f(x_k)\right]\Delta x$ = (2+5+10+17)*1.0 =34.0

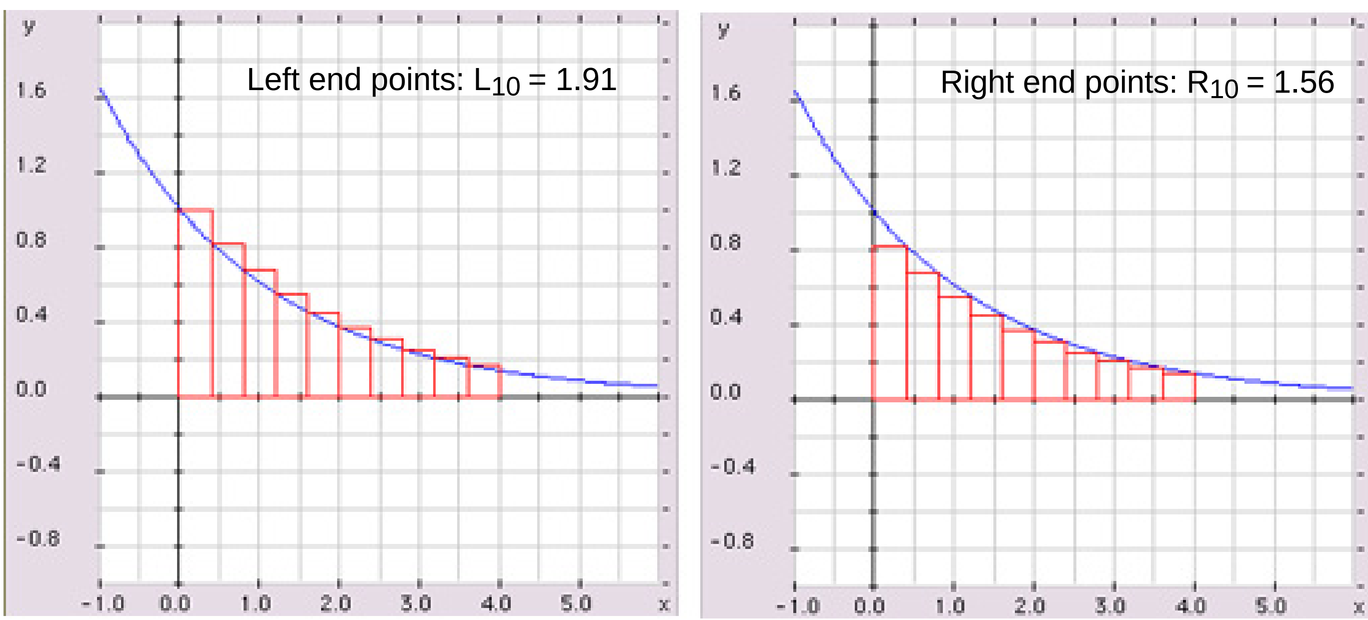
These are called Riemann sums --
Denoted L_4 for left end-points with 4 intervals
and R_4 for right end-points with 4 intervals

Question: Suppose you're given a specific problem and you are required to use exactly 10 rectangles. Are there good & bad choices for how exactly to pick the rectangles?

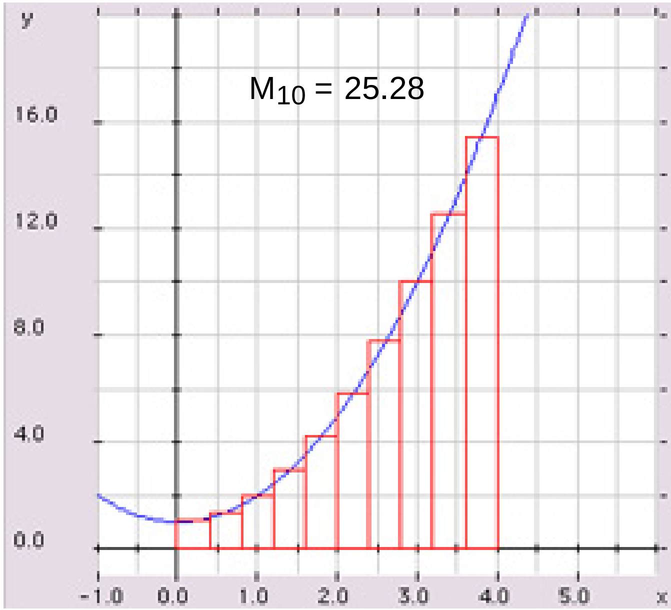
Example 1: $y = f(x) = x^2 + 1$, from x=0 to x=4, with 10 equal intervals.



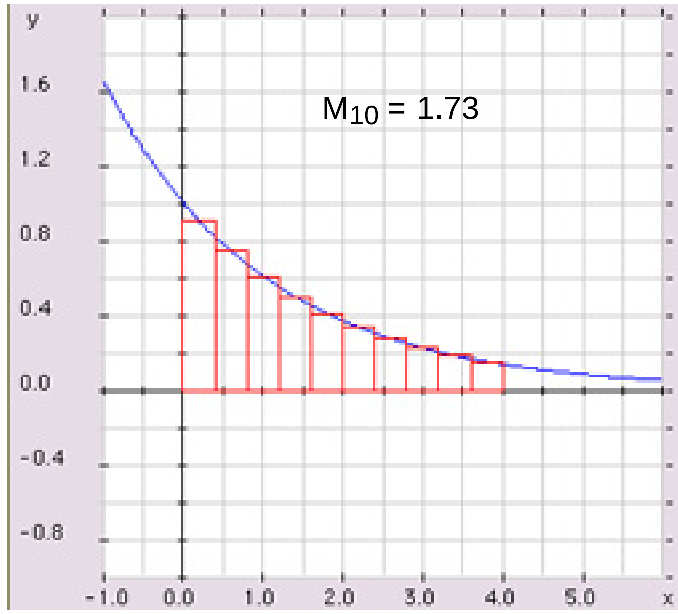
Example 2: $y = f(x) = e^{-x/2}$ from x=0 to x=4, with 10 equal intervals.



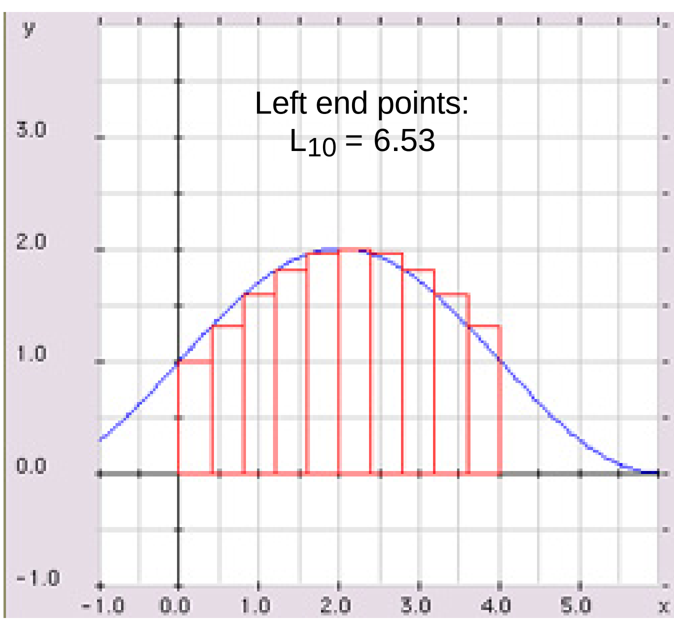
Example 1 with mid-points

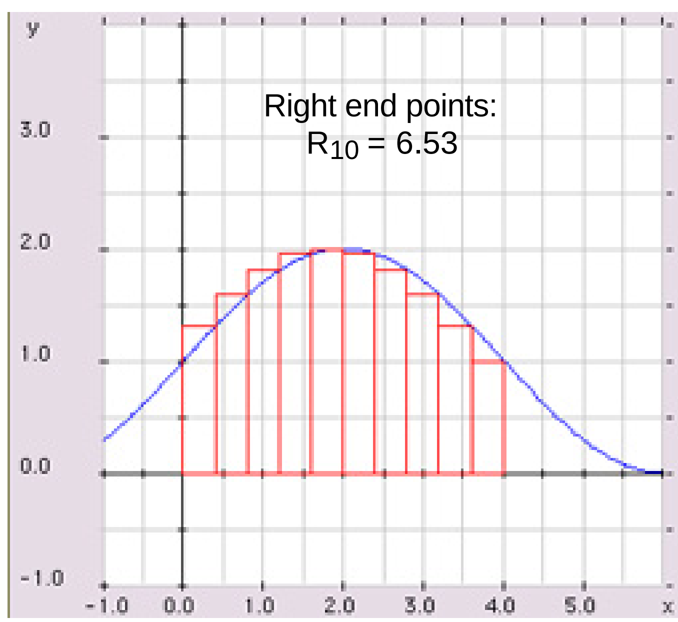


Example 2 with mid-points



Example 3: $y = f(x) = \sin\left(\frac{\pi}{4}x\right) + 1$ from x=0 to x=4, with 10 equal intervals.





Conclusions

(I) If f(x) is monotonically increasing from a to b then:

The left end-points underestimate the true area.

The right end-points overestimate the true area.

(II) If f(x) is monotonically decreasing from a to b then:

The left end-points overestimate the true area.

The right end-points underestimate the true area.

(III) If f(x) is not monotonic on a to b then each type of Riemann sum (left & right) may underestimate or overestimate the true area -- we cannot predict without knowing the exact form of f(x).

The exact area under a curve

Key points:

* The approximate area, clearly, gets more & more accurate as the number of rectangles (say, *n*) increases.

* It follows that the exact area is obtained when *n* goes to ∞

* In general, the approximate area with *n* rectangles (of equal width) has the form

$$A_n = \left[\sum_{k=0}^{n-1} f(x_k) \right] \Delta x \qquad \text{OR} \qquad \overline{A}_n = \left[\sum_{k=1}^{n} f(x_k) \right] \Delta x$$

* Thus, the exact area is:

$$A = \lim_{n \to \infty} \left[\sum_{k=0}^{n-1} f(x_k) \right] \Delta x = \lim_{n \to \infty} \left[\sum_{k=1}^{n} f(x_k) \right] \Delta x$$

* This comprises the fundamental definition of the definite integral:

Definite integral = Limit of a Riemann sum as the summation goes to ∞

Notation: $\int_a^b f(x)\, dx = \lim_{n \to \infty} \left[\sum_{k=1}^{n} f(x_k^*) \right] \Delta x \qquad \left\{ \text{where } \Delta x = \frac{(b-a)}{n} \right\}$

Here x_k^* denotes any point between x_{k-1} and x_k.

Key point to note: The definite integral of a function is a number, NOT a function.

Recall, the indefinite integral of a function is a function.

Definite integral

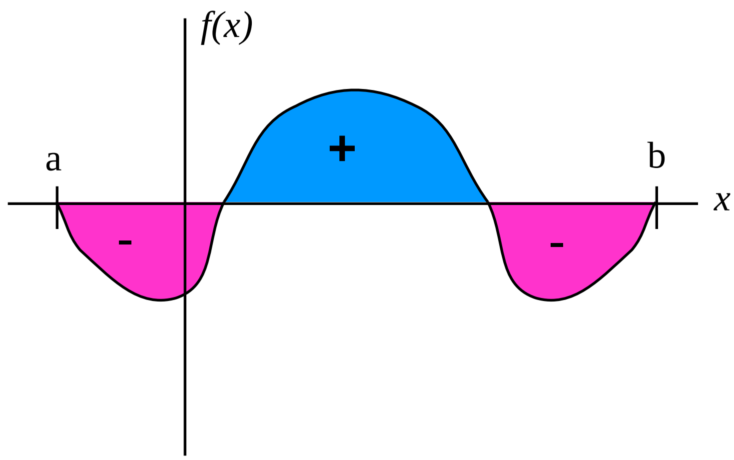
Geometric Interpretation:

For any $f(x)$ continuous (and positive) on the interval [a,b], the definite integral $\int_a^b f(x)\,dx$ is the (exact) area under the graph of $f(x)$ from a to b.

Question: What if $f(x)$ is not positive?

A: Easy to figure out using the Riemann sum interpretation.

If $A_n = \left[\sum_{k=1}^{n} f(x_k^*)\right]\Delta x$, what happens when $f(x_k^*)$ is negative?



Regions where f is positive contribute positive values to the definite integral.
Regions where f is negaitive contribute negaitive values.

The net value of $\int_a^b f(x)\,dx$ is obtained by adding the positive regions and subtracting the negaitive regions.

Proof strategy for evaluation theorem

Prelude: How can one prove equality between any two entities, i.e., LHS=RHS?

One way: Start with LHS (or RHS) -- then manipulate it and transform it into the other.

E.g., Prove that $(x+1)^2 = x^2 + 2x + 1$

For the evaluation theorem, we must prove: $\int_a^b f(x)\, dx = F(b) - F(a)$

where $F(x)$ is a function that satisfies $\frac{dF}{dx} = f$.

Proof strategy:

(1) Start with RHS, which is $F(b)$ - $F(a)$.

(2) Transform this into a summation involving n terms over the interval $[a, b]$.

E.g., $F(b) - F(a) = F(b) + [(p-p) + (q-q) + (r-r) + \ldots] - F(a)$

Select $p = F(x_{n-1})$, $q = F(x_{n-2})$, $r = F(x_{n-3})$, $\ldots$; with $b = x_n, a = x_0$

This gives: $F(b) - F(a) = F(x_n) - F(x_{n-1}) + F(x_{n-1}) - F(x_{n-2}) + \ldots + F(x_1) - F(x_0)$

$$\therefore \quad F(b) - F(a) = \sum_{k=1}^{n} \left[F(x_k) - F(x_{k-1}) \right]$$

(Now we need to figure out a way to bring f into the above summation.)

(3) Use mean value theorem to relate differences in F to its derivative $\frac{dF}{dx}$.

(4) Use the condition that F is required to satisfy: $\frac{dF}{dx} = f$, to bring f into the picture. Take limits as $n \to \infty$, and we obtain the required LHS. Hence we have shown that LHS=RHS.

[**Aside:** What does the mean value theorem really say?]