Definite Integrals

Definition: A **partition** $\mathcal{P}$ of the closed interval $I = [a, b]$ is a set of closed intervals
$I_j = [x_{j-1}, x_j]$ defined by a sequence of numbers

$$x_0 < x_1 < x_2 \ldots < x_{n-1} < x_n$$

where $a = x_0$ and $x_n = b$. We denote the length $x_j - x_{j-1}$ of I_n by Δx_j. The numbers x_j are called **partition points**.

Example: If $a = 1, b = 10$, and we let $x_0 = a = 1$, $x_1 = 2$, $x_3 = 5$, $x_4 = 9$, and $x_5 = 10 = b$, then the partition is $\mathcal{P} = \{[1, 2], [2, 5], [5, 9], [9, 10]\} = \{I_1, I_2, I_3, I_4\}$, where $I_1 = [1, 2]$, $I_2 = [2, 5]$, $I_3 = [5, 9]$, $I_4 = [9, 10]$, and $\Delta x_1 = 2 - 1 = 1$, $\Delta x_2 = 5 - 2 = 3$, $\Delta x_3 = 9 - 5 = 4$, $\Delta x_4 = 10 - 9 = 1$.

Definition: A **tagset** τ for the partition $\mathcal{P}$ is a set of numbers $\tau = \{t_1 \leq t_2 \leq \cdots \leq t_n\}$ which satisfy $t_j \in I_n$.

Example: In the previous example, let $t_1 = 1$, $t_2 = 3$, $t_3 = 7.5$, and $t_4 = 10$, so that $\tau = \{1, 3, 7.5, 10\}$.

Definition: $\mathcal{F}([a,b])$ is defined to be the set of all functions with domain $[a,b]$.

Definition: If $f \in \mathcal{F}([a,b])$, and $\mathcal{P}$ is a partition of $[a,b]$ with tagset $\mathcal{T}$, then the **Riemann sum** of f with partition P and tagset T is defined to be

$$\mathcal{S}(f, \mathsf{P}, \mathsf{T}) = f(t_1)\Delta x_1 + f(t_2)\Delta x_2 + \cdots + f(t_n)\Delta x_n$$

Example: Using the previous example,

$\mathcal{S}(f, \mathsf{P}, \mathsf{T}) = \mathcal{S}(f, \{[1,2],[2,5],[5,9],[9,10]\}, \{1,3,7.5,10\}) =$

$f(t_1)\Delta x_1 + f(t_2)\Delta x_2 + f(t_3)\Delta x_3 + f(t_4)\Delta x_4 =$

$f(1)(1) + f(3)(3) + f(7.5)(4) + f(10)(1) =$

$f(1) + 3f(3) + 4f(7.5) + f(10)$

Examples of special Riemann sums:

(1) If we let $t_j = x_{j-1}$, we get the *Left-hand sum* of f associated with the partition P:
$\mathcal{L}(f, \mathsf{P}) = f(x_0)\Delta x_1 + f(x_1)\Delta x_2 + \cdots + f(x_{n-1})\Delta x_n$

(2) If we let $t_j = x_j$, we get the *Right-hand sum* of f associated with the partition P:
$\mathcal{R}(f, \mathsf{P}) = f(x_1)\Delta x_1 + f(x_2)\Delta x_2 + \cdots + f(x_n)\Delta x_n$

(3) If we let $t_j = (x_{j-1} + x_j)/2$, we get the *Midpoint sum* of f associated with the partition P:
$\mathcal{M}(f, \mathsf{P}) = f(\frac{x_0+x_1}{2})\Delta x_1 + f(\frac{x_1+x_2}{2})\Delta x_2 + \cdots + f(\frac{x_{n-1}+x_n}{2})\Delta x_n$

Another natural estimate of the area under a curve is the *Trapezoidal Estimate* formed by adding up the areas of the trapezoids with base $[x_{j-1}, x_j]$ and heights $f(x_{j-1})$ and $f(x_j)$:

$$\mathcal{T}(f, \mathsf{P}) = \frac{f(x_0) + f(x_1)}{2}\Delta x_1 + \frac{f(x_1) + f(x_2)}{2}\Delta x_2 + \cdots + \frac{f(x_{n-1}) + f(x_n)}{2}\Delta x_n.$$

It is easily verified that $\mathcal{T}(f, \mathsf{P}) = \dfrac{\mathcal{L}(f, \mathsf{P}) + \mathcal{R}(f, \mathsf{P})}{2}$, that is, that it is the average of the Left-hand and Right-hand sums. It is not itself a Riemann sum.

If we now insist that all the functions f are continuous on the closed interval we can define two more very important Riemann sums:

If s_j is selected so that the minimum m_i of $f(x)$ on the interval I_j occurs at s_j, we get the *Inscribed sum* of f associated with the partition P:

$$\mathcal{I}(f, \mathsf{P}) = f(s_1)\Delta x_1 + f(s_2)\Delta x_2 + \cdots + f(s_n)\Delta x_n$$
$$= m_1\Delta x_1 + m_2\Delta x_2 + \cdots + m_n\Delta x_n$$

If $f(x) \geq 0$ on $[a, b]$, this is of course the sum of the areas of the rectangles inscribed under the graph of $y = f(x)$ with the intervals I_j as their bases. Note that the numbers s_j need not be unique, since a function can attain its minimum at more than one point in an interval. This does not affect the value of the Inscribed sum. It was necessary to assume that f was continous so as to be certain that it would actually have a minimum value on every interval I_j.

If t_j is selected so that the maximum M_i of $f(x)$ on the interval I_j occurs at t_j, we get the *Exscribed sum* of f associated with the partition P:
$\mathcal{E}(f, \mathsf{P}) = f(t_1)\Delta x_1 + f(t_2)\Delta x_2 + \cdots + f(t_n)\Delta x_n$
$= M_1\Delta x_1 + M_2\Delta x_2 + \cdots + M_n\Delta x_n$

If $f(x) \geq 0$ on $[a, b]$, this is of course the sum of the areas of the rectangles exscribed over the graph of $y = f(x)$ with the intervals I_j as their bases. Note again that the numbers t_j need not be unique.

It is clear that if f is continuous on $[a, b]$, then for any tagset T we have:

$$\mathcal{I}(f, \mathsf{P}) \leq \mathcal{S}(f, \mathsf{P}, \mathsf{T}) \leq \mathcal{E}(f, \mathsf{P})$$

Definition: The *mesh* or *norm* of the partition P is the number

$$||P|| = \max\{\Delta x_j : 1 \leq j \leq n\}.$$

Lemma: If f has a continuous derivative on the interval $I_j = [x_{j-1}, x_j]$, then, by the Mean Value Theorem,

$M_j - m_j = f(t_j) - f(s_j) = f'(c_j)(t_j - s_j)$ for some c_j between s_j and t_j. Thus $M_j - m_j = |f'(c_j)||t_j - s_j| \leq |f'(c_j)|\Delta x_j$.

Theorem: If f is has a continous derivative on $[a,b]$, let $M = \max\{|f'(x)| : x \in [a,b]\}$. Then $\mathcal{E}(f,\mathsf{P}) - \mathcal{I}(f,\mathsf{P}) \le M(b-a)||\mathsf{P}||$.

Proof: $\mathcal{E}(f,\mathsf{P}) - \mathcal{I}(f,\mathsf{P}) =$

$(M_1 - m_1)\Delta x_1 + (M_2 - m_2)\Delta x_2 + \cdots + (M_n - m_n)\Delta x_n =$

$= |f'(c_1)||t_1 - s_1|\Delta x_1 + |f'(c_2)||t_2 - s_2|\Delta x_2 + \cdots + |f'(c_n)||t_n - s_n|\Delta x_n$

$\le M||\mathsf{P}||\Delta x_1 + M||\mathsf{P}||\Delta x_2 + \cdots + M||\mathsf{P}||\Delta x_n$

$\le M||\mathsf{P}||(\Delta x_1 + \cdots + \Delta x_n)$

$= M(b-a)||\mathsf{P}||$ **Q.E.D.**

Corollary: If f has a continuous derivative on $[a,b]$, then for any partition P and tagset T we have:

$\mathcal{E}(f,\mathsf{P}) - \mathcal{R}(f,\mathsf{P},\mathsf{T}) \le M(b-a)||\mathsf{P}||$ and $\mathcal{R}(f,\mathsf{P},\mathsf{T}) - \mathcal{I}(f,\mathsf{P}) \le M(b-a)||\mathsf{P}||$

Thus at this point we can begin to be convinced that $\lim_{||\mathsf{P}||\to 0} S(f, \mathsf{P}, \mathsf{T})$ should exist if f has a continuous first derivative. As a matter of fact, it can be shown that this limit exists if f is just continuous on $[a, b]$, but this is much more difficult to prove.

Definition: If $\lim_{||\mathsf{P}||\to 0} S(f, \mathsf{P}, \mathsf{T})$ exists, it is called the **definite integral** of f from a to b and is denoted by $\int_a^b f(x)dx$

Properties of the Definite Integral

$$\int_a^b h\,dx = h(b-a)$$

— if $h > 0$ this is the area of the rectangle with base $[a,b]$ and height h.

$$\int_a^b \left(f(x) + g(x)\right) dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

$$\int_a^b \left(f(x) - g(x)\right) dx = \int_a^b f(x)dx - \int_a^b g(x)dx$$

$\displaystyle\int_a^b kf(x)dx = k\int_a^b f(x)dx$, where k is a constant.

If $a > b$ we define $\displaystyle\int_a^b f(x)dx = -\int_b^a f(x)dx$.

Then we have $\displaystyle\int_a^b kf(x)dx = \int_a^c kf(x)dx + \int_c^b kf(x)dx$
for any numbers a, b, c for which all three definite integrals are defined.

Clearly we have $\int_a^a f(x)dx = 0$.

Inequalities of Definite Integrals

If $f(x) \geq 0$ on $[a, b]$ and $a < b$ then $\int_a^b f(x)dx \geq 0$.

Indeed, if f is also continuous at a number c in $[a, b]$ and $f(c) > 0$, we can say that
$\int_a^b f(x)dx > 0$.

We can use these facts to derive some more general inequalities: Suppose $f(x) \geq g(x)$ on $[a, b]$.

Then $f(x) - g(x) \geq 0$ on $[a, b]$, so $\int_a^b \left(f(x) - g(x)\right) dx \geq 0$,

but we also have $\int_a^b \left(f(x) - g(x)\right) dx = \int_a^b f(x)dx - \int_a^b g(x)dx$

so $\int_a^b f(x)dx - \int_a^b g(x)dx \geq 0$, and therefore $\int_a^b f(x)dx \geq \int_a^b g(x)dx$

Thus $f(x) \geq g(x)$ on $[a, b] \quad \Longrightarrow \int_a^b f(x)dx \geq \int_a^b g(x)dx$

We can use this to get another important chain of inequalities: if $m \leq f(x) \leq M$ on $[a, b]$,

then $\int_a^b m dx \leq \int_a^b f(x) dx$ and $\int_a^b f(x) dx \leq \int_a^b M dx$

But $\int_a^b m dx = m(b - a)$ and $\int_a^b M dx = M(b - a)$, so we have

$$m(b - a) \leq \int_a^b f(x) dx \leq M(b - a).$$

Dividing every term in this chain of inequalities by the positive number $b - a$ results in

$$m \leq \frac{\int_a^b f(x) dx}{b - a} \leq M.$$

The expression $\frac{\int_a^b f(x) dx}{b - a}$ is called the **average value** of $f(x)$ on the interval $[a, b]$, and it should come as no surprise that it lies between the minimum and maximimum values of $f(x)$ on $[a, b]$.

Mean Value Theorem for Definite Integrals

If f is continuous on $[a, b]$, and we let $m = \min\{f(x) | a \leq x \leq b\}$, and $m = \min\{f(x) | a \leq x \leq b\}$, then we know by the Intermediate Value Theorem that

$f(x)$ takes on all values between m and M, so in particular it takes on the *average value* for some number c in $[a, b]$. We therefore have

$f(c) = \dfrac{\int_a^b f(x)dx}{b-a}$, or,

in the usual form in which the Mean Value Theorem for Definite Integrals is stated:

$$\int_a^b f(x)dx = f(c)(b-a)$$

Geometric and Algebraic Areas .

The definite integral $\int_a^b f(x)dx$ is called the **algebraic** area between the graph of $y = f(x)$ and the x-axis. It is the sum of all of the areas of regions where f is positive *minus* the sum of the all the areas of regions where f is negative.

The definite integral $\int_a^b |f(x)|dx$ is called the **geometric** area between the graph of $y = f(x)$ and the x-axis. It is the sum of all of the areas of regions where f is positive *plus* the sum of the all the areas of regions where f is negative.

Clearly $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$

If f never changes sign on $[a, b]$ then the two quantities will be equal.

If f changes sign on $[a, b]$ and is continuous, then the two quantities will *not* be equal.

Notation

(1) We define $\mathcal{D}(f) = f'$, the derivative of f. $\mathcal{D}$ is usually called the differential operator.

(2) We define $\mathcal{I}(f) = F$, where $F(x) = \int_a^x f(t)dt$, the desired antiderivative F of f for which $F(a) = 0$. $\mathcal{I}$ is usually called an integral operator.

Using this notation, we state, without proof, the:

Fundamental Theorem of Calculus:

If f is continous on $[a, b]$, then $\mathcal{D}(\mathcal{I}(f)) = f$, and if F is differentiable on $[a, b]$ then $\mathcal{I}(\mathcal{D}(F))(b) = \displaystyle\int_a^b \mathcal{D}(F)dx = F(b) - F(a)$.

We *will* prove this with the additional assumption that f has a continuous first derivative.

Notation: instead of writing $F(b) - F(a)$, it will be convenient to use instead the expression

$$F(x)|_a^b,$$

(or even $F(x)|_{x=a}^{x=b}$ when we want to be clear about what is the variable of integration.)

Thus $x^2\,\big|_2^5 = 5^2 - 2^2 - 25 - 4 = 21$

Example of an area calculation: suppose we wish to find the area between the x-axis and $y = 3x^2$ from $x = 4$ to $x = 8$.

All we need to do is find an antiderivative $F(x)$ of $3x^2$ — $F(x) = x^3$ will do nicely, and compute

$F(8) - F(4) = 8^3 - 4^3 = 512 - 64 = 448$

The notation generally used looks like this:

$$\int_4^8 3x^2 dx = x^3\,\Big|_4^8 = 8^3 - 4^3 = 512 - 64 = 448$$

Some Handy Antiderivative, or Indefinite Integral, Formulas:

$$\int (f(x) + g(x))\, dx = \int f(x)dx + \int g(x)dx$$

$$\int (f(x) - g(x))\, dx = \int f(x)dx - \int g(x)dx$$

$\int kf(x)dx = k \int f(x)dx$, where k is a constant.

$\int x^n dx = \dfrac{x^{n+1}}{n+1} + C$, if $n \neq -1$

$$\int \sin x\, dx = -\cos x + C \qquad \int \cos x\, dx = \sin x + C$$

$$\int \sec^2 x\, dx = \tan x + C \qquad \int \sec x \tan x\, dx = \sec x + C$$

Definition: The *regular* partition of $[a, b]$ with n intervals is formed by letting

$$\Delta x = \frac{b-a}{n}, \; x_j = a + j\Delta x. \text{ We have } ||\mathsf{P}|| = \Delta x.$$

For regular partitions we get nicer formulas for the special Riemann sums we have defined above:

(1) The *Left-hand sum* of f becomes

$$\mathcal{L}(f, \mathsf{P}) = \mathcal{L}_n(f) = \sum_{i=1}^{n} f(x_{i-1})\Delta x = [f(x_0) + f(x_1) + \cdots + f(x_{n-1}]\,\Delta x =$$

$$\left[\sum_{i=1}^{n} f\left(a + (i-1)\frac{b-a}{n}\right)\right]\frac{b-a}{n}$$

(2) The *Right-hand sum* of f becomes

$$\mathcal{R}(f,\mathsf{P}) = \mathcal{R}_n(f) = \sum_{i=1}^{n} f(x_i)\Delta x = [f(x_1) + f(x_2) + \cdots + f(x_n)]\,\Delta x =$$

$$\left[\sum_{i=1}^{n} f\left(a + i\frac{b-a}{n}\right)\right]\frac{b-a}{n}$$

---

(3) The *Midpoint sum* of f becomes

$$\mathcal{M}(f,\mathsf{P}) = \mathcal{M}_n(f) = \sum_{i=1}^{n} f\left(\frac{x_{i-1} + f(x_i)}{2}\right)\Delta x =$$

$$\left[f\left(\frac{x_0 + x_1}{2}\right) + f\left(\frac{x_1 + x_2}{2}\right) + \cdots + f\left(\frac{x_{n-1} + x_n}{2}\right)\right]\Delta x =$$

$$\left[\sum_{i=1}^{n} f\left(a + \left(i - \frac{1}{2}\right)\frac{b-a}{n}\right)\right]\frac{b-a}{n}$$

---

(4) The *Inscribed sum* of f becomes $\mathcal{I}(f, \mathsf{P}) = \mathcal{I}_n(f) = [m_1 + m_2 + \cdots + m_n]\Delta x$

(5) The *Exscribed sum* of f becomes $\mathcal{E}(f, \mathsf{P}) = \mathcal{E}_n(f) = [M_1 + M_2 + \cdots + M_n]\Delta x$

The *Trapezoidal Estimate* can now be written as $\mathcal{T}(f, \mathsf{P}) = \mathcal{T}_n(f) = \frac{1}{2}(f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_n - 1) + f(x_n))\Delta x =$

$$\left[f(a) + 2\sum_{i=1}^{n-1} f\left(a + i\frac{b-a}{n}\right) + f(b) \right] \frac{b-a}{n}$$

Theorem If f has continuous derivative on $[a, b]$,

then $\mathcal{E}_n(f) - \mathcal{I}_n(f) \leq M\frac{(b-a)^2}{n}$, so

$$\lim_{n\to\infty} \mathcal{E}_n(f) - \mathcal{I}_n(f) = 0$$

In plain English, the difference between the sums of exscribed and inscribed rectangles goes to 0 as the number of rectangles approaches infinity. This means that the limit as $n \to \infty$ of any of our sums is equal to the desired value

$$\int_a^b f(x)dx$$

Example:

Let $y = f(x) = x^2$, $a = 0$, $b = 1$. Then $x_i = \frac{i}{n}$

(1) $\mathcal{L}_n(f) = \sum_{i=0}^{n-1} \left(\frac{i}{n}\right)^2 \frac{1}{n} = \frac{1}{n^3} \sum_{i=1}^{n-1} i^2 =$

$$\frac{1}{n^3}\frac{(n-1)(n)(2n-1)}{6} = \frac{1}{n^2}\frac{(n-1)(2n-1)}{6} = \left(1 - \frac{1}{n}\right)\left(\frac{1}{3} - \frac{1}{6n}\right) =$$

$$\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}$$

(2) $\mathcal{R}_n(f) = \sum_{i=1}^{n} \left(\frac{i}{n}\right)^2 \frac{1}{n} = \frac{1}{n^3} \sum_{i=1}^{n} i^2 = \frac{1}{n^3} \frac{n(n+1)(2n+1)}{6} =$

$$\frac{1}{n^2} \frac{(n+1)(2n+1)}{6} = \left(1 + \frac{1}{n}\right)\left(\frac{1}{3} + \frac{1}{6n}\right) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}$$

Now $\mathcal{T}_n(f) = \frac{\mathcal{L}_n(f) + \mathcal{R}_n(f)}{2} = \frac{1}{3} + \frac{1}{6n^2}$

$$(3)\ \mathcal{M}_n(f) = \sum_{i=1}^{n} \left(\frac{2i-1}{2n}\right)^2 \frac{1}{n} = \frac{1}{4n^3} \sum_{i=1}^{n} (2i-1)^2 =$$

$$\frac{1}{4n^3} \left(\sum_{i=1}^{2n} i^2 - \sum_{i=1}^{n} (2i)^2 \right) = \frac{1}{4n^3} \left(\sum_{i=1}^{2n} i^2 - 4 \sum_{i=1}^{n} i^2 \right) =$$

$$\frac{1}{4n^3} \left(\frac{2n(2n+1)(2(2n)+1)}{6} - 4\frac{n(n+1)(2n+1)}{6} \right) =$$

$$\frac{n(2n+1)}{12n^3} \left((2(2n)+1) - 2(n+1) \right) = \frac{2n+1}{12n^2} (4n+1-2n-2) =$$

$$\frac{2n+1}{12n^2} (2n-1) = \frac{4n^2-1}{12n^2} = \frac{1}{3} - \frac{1}{12n^2}$$

---

(4) $\mathcal{I}_n(f) = \mathcal{L}_n(f) = \frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}$

(5) $\mathcal{E}_n(f) = \mathcal{R}_n(f) = \frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}$

We note that $\mathcal{E}_n(f) - \mathcal{I}_n(f) = \left(\frac{1}{3} + \frac{1}{2n} + \frac{1}{6n^2}\right) - \left(\frac{1}{3} - \frac{1}{2n} + \frac{1}{6n^2}\right) = \frac{1}{n}$

Note that all of these sums go to $\frac{1}{3}$ as $n \to \infty$.

Of course, the easy way to compute $\int_0^1 x^2 dx$ is to use the the Fundamental Theorem of Calculus and the fact that an antiderivative of x^2 is $F(x) = \frac{x^3}{3}$ and simply evaluate

$$F(1) - F(0) = \frac{1^3}{3} - \frac{0^3}{3} = \frac{1}{3}$$