Math 8	Instructor: Padraic Bartlett

Lecture 6: Integration and the Mean Value Theorem

Week 6 | *Caltech 2012*

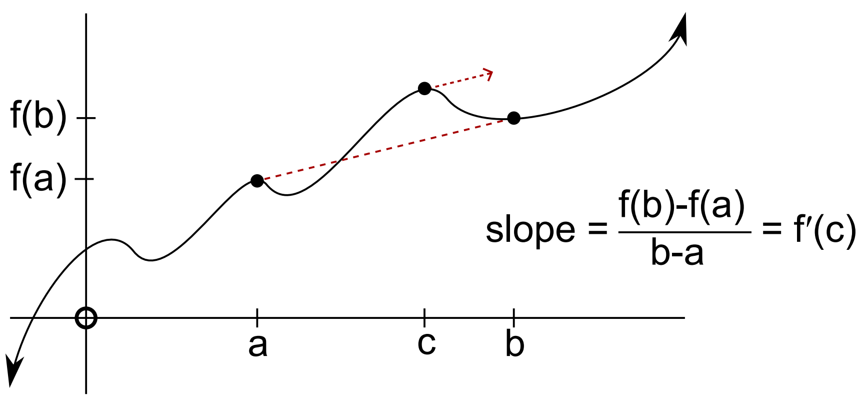
1 The Mean Value Theorem

The Mean Value Theorem (abbreviated MVT) is the following result:

Theorem. Suppose that f is a continuous function on the interval $[a, b]$ that's differentiable on (a, b). Then there is some value c such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

In other words, there is some point c between $f(a)$ and $f(b)$ such that the derivative at that point is equal to the slope of the line segment connecting $(a, f(a))$ and $(b, f(b))$. The following picture illustrates this claim:



A common/practical example of the mean value theorem's use in day-to-day life is on toll roads: suppose that you get on a road at 2pm, travel 150 miles, and get off the road at 4pm. The mean value theorem says that at some point in time, you must have traveled 75mph, because in order to travel 150 miles in 2 hours your **average** speed must have been 75 mph[1].

The mean value theorem, much like the intermediate value theorem, is usually not a tough theorem to understand: the tricky thing is realizing when you should try to **use** it. Roughly speaking, you want to use the mean value theorem **whenever you want to turn information about a function into information about its derivative, or vice-versa**. We work two examples of how the mean value theorem is used below:

[1] Toll booths in some states will actually use this information to automatically print out speeding tickets for people! Because the ticket you purchace when you enter the toll road states when and where you entered the highway, and you must show your ticket to leave the highway, they know when and where you entered and left. This therefore determines your average speed, which (by the mean value theorem) you must have attained at some point in your journey!

1.1 Applications of the Mean Value Theorem

Example. Consider the equation

$$(x+y)^n = x^n + y^n.$$

If either x or y are zero, this equation holds; as well, if $x = -y$ and n is odd, this equation also holds. Are there any other values of $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$ that are solutions of this equation?

Proof. First, notice the following lemma we can prove using the mean value theorem:

Lemma 1. *Suppose you have a differentiable function f with k distinct roots $a_1 < a_2 < \ldots a_k$. Then f' has at least $k-1$ distinct roots $b_1 < b_2 < \ldots b_{k-1}$, such that*

$$a_1 < b_1 < a_2 < b_2 < \ldots b_{k-1} < a_k.$$

Proof. We know f is differentiable and continuous on $[a_1, a_2]$: therefore, by the mean value theorem, we can find some value b_1 such that

$$f'(b_1) = \frac{f(a_2) - f(a_1)}{a_2 - a_1}.$$

But we know that $f(a_2) = f(a_1) = 0$, because a_1 and a_2 are roots of f: therefore, we actually have that

$$f'(b_1) = \frac{f(a_2) - f(a_1)}{a_2 - a_1} = \frac{0-0}{a_2 - a_1} = 0,$$

and that b_1 is a root of f'. Repeating this for all of the other pairs a_j, a_{j+1} will create the $k-1$ roots $b_1, \ldots b_{k-1}$ that we claimed exist. $\square$

What does this lemma tell us about our problem? Well: pick any fixed nonzero value of y in $\mathbb{R}$, and look at the equation

$$f(x) = (x+y)^n - x^n - y^n.$$

We are currently trying to find out which values of x give us a root of this equation. Currently, we know the following roots:

- If n is even, the only root we know is $x = 0$.
- If n is odd, we know that $x = 0, x = -y$ are two roots.

Can there be any more distinct roots? Well: by our earlier lemma, we know that if our function were to have k roots, its derivative would have to have at least $k-1$ roots. So: how many roots can f' have?

Calculating tells us that

$$f'(x) = n(x+y)^{n-1} - nx^{n-1},$$

which is equal to 0 whenever

$$\begin{aligned} 0 &= n(x+y)^{n-1} - nx^{n-1} \\ \Leftrightarrow nx^{n-1} &= n(x+y)^{n-1} \\ \Leftrightarrow x^{n-1} &= (x+y)^{n-1}. \end{aligned}$$

Here, we have two cases. If n is even, we know that $n-1$ is odd, and therefore the equation $x^{n-1} = (x+y)^{n-1}$ is equivalent to the claim

$$\begin{aligned} x &= (x+y) \\ \Leftrightarrow 0 &= y, \end{aligned}$$

which contradicts our nonzero choice of y. So when n is even, we cannot have that f' has a root: by our lemma, this means that when n is even, f cannot have more than 1 distinct root. So the only root of f when n is even is $x = 0$, because we've just shown that there can be no others.

If n is odd, we know that $n-1$ is even, in which case our equation $x^{n-1} = (x+y)^{n-1}$ is equivalent to the claim

$$\begin{aligned} |x| &= |x+y| \\ \Leftrightarrow \pm x &= x+y \\ \Leftrightarrow y &= 0 \text{ or } y = -2x. \end{aligned}$$

So, in the case where n is odd and y is nonzero, f' has exactly one root at $x = -\frac{y}{2}$. By our lemma, this means that when n is odd, f cannot have more than 2 distinct roots. So the only roots of f when n is odd are $x = 0$ and $x = -y$, because we've just shown that there can be no others. □

The example above showed how we could turn information about our function (specifically, knowledge of where its roots are) into information about the derivative. The mean value theorem can also be used to turn information about the derivative into information about the function, as we illustrate here:

Example. Let $p(t)$ denote the current location of a particle moving in a one-dimensional space. Suppose that $p(0) = 0, p(1) = 1$, and $p'(0) = p'(1) = 0$. Show that there must be some point in time in $[0, 1]$ where $|p''(t)| \geq 4$.

Proof. We proceed by contradiction: i.e. suppose instead that for every $t \in [0, 1]$ we have $|p''(t)| < 4$.

What can we do from here? Well: we have some boundary conditions ($p(0) = 0, p(1) = 1, p'(0) = 0, p'(1) = 0$) and one global piece of information ($|p''(t)| < 4$). How can we turn this knowledge of the second derivative into information about rest of the function?

Well: if we apply the mean value theorem to the function $p'(t)$, what does it say? It tells us that on any interval $[a, b]$, there is some $c \in (a, b)$ such that

$$\frac{p'(b) - p'(a)}{b - a} = (p')'(x) = p''(c).$$

In other words, it relates the first and second derivatives to each other! So, if we apply our known bound $|p''(t)| < 4, \forall t \in [0, 1]$, we've just shown that

$$\left| \frac{p'(b) - p'(a)}{b - a} \right| = |p''(c)| < 4,$$

for any $a < b \in [0, 1]$. In particular, if we set $a = 0, b = t$ and remember our boundary condition $p'(0) = 0$, we've proven that

$$\begin{aligned} &\left| \frac{p'(t) - p'(0)}{t - 0} \right| = \left| \frac{p'(t) - 0}{t} \right| = \frac{|p'(t)|}{t} < 4 \\ \Rightarrow &|p'(t)| < 4t. \end{aligned}$$

Similarly, if we let $a = 1 - t$ and $b = 1$, we get

$$\begin{aligned} &\left| \frac{p'(1) - p'(1-t)}{1 - (1-t)} \right| = \left| \frac{0 - p'(1-t)}{t} \right| = \frac{|p'(1-t)|}{t} < 4 \\ \Rightarrow &|p'(1-t)| < 4t. \end{aligned}$$

Excellent! We've turned information about the second derivative into information about the first derivative.

Pretend, for the moment, that you're back in your high school calculus courses, and you know how to find antiderivatives. In this situation, you've got a function $p(t)$ with the following properties:

- $p(0) = 0$,
- $p(1) = 1$,
- $|p'(t)| < 4t$, and
- $|p'(1-t)| < 4t$.

What do you know about $p(t)$? Well: if $p'(t) < 4t$, you can integrate to get that $p(t) < 2t^2 + C$, for some constant C: using our boundary condition $p(0) = 0$ tells us that in specific we can pick $C = 0$, and we have

$$p(t) < 2t^2, \forall t \in (0, 1).$$

Similarly, if we use our observation that $-p(1-t) > -4t$, we can integrate to get that $p(1-t) > -2t^2 + C$: using our boundary condition $p(1) = 1$ tells us that in specific we can pick $C = 1$, which gives us

$$p(1-t) > -2t^2 + 1, \forall t \in (0,1).$$

But what happens if we plug in $t = \frac{1}{2}$? In our first bound, we have $p\left(\frac{1}{2}\right) < 2\left(\frac{1}{2}\right)^2 = \frac{1}{2}$. Conversely, in our second bound we have $p\left(1 - \frac{1}{2}\right) > -2\left(\frac{1}{2}\right)^2 + 1 = \frac{1}{2}$: in other words, at $\frac{1}{2}$ our function must both be greater than and less than $\frac{1}{2}$! This is clearly impossible, so we've reached a contradiction …

Assuming, of course, that we know how to perform antidifferentiation. Which we don't (at least, not officially!) How can we solve this problem using just the mean value theorem?

Earlier, we turned information about the second derivative into information about the first derivative. Can we do that trick again to get information about the original function? Well: let's try applying the mean value theorem to the function f, on the interval $[0, t]$. This tells us that there is some value of $c \in (0, t)$ such that

$$\frac{f(t) - f(0)}{t - 0} = f'(c).$$

If we use our boundary condition $f(0) = 0$ and our bound $|f'(c)| < 4c < 4t, \forall c \in (0, t)$, this becomes

$$\left|\frac{f(t) - 0}{t}\right| = |f'(c)| < 4t \Rightarrow \qquad |f(t)| < 4t^2.$$

This is …not the bound of $2t^2$ that we got by antidifferentiating. What can we do here? Well: what happens if we take this bound, and look at the interval $[t, 2t]$? Applying the mean value theorem there tells us that there is some $c \in (t, 2t)$ such that

$$\begin{aligned}
&\frac{f(2t) - f(t)}{2t - t} = f'(c) \\
\Rightarrow &|f(2t) - f(t)| = t|f'(c)| < t \cdot 4c < t \cdot 4 \cdot 2t = 8t^2 \\
\Rightarrow &|f(2t)| < 8t^2 + |f'(t)| < 4t^2 + 8t^2.
\end{aligned}$$

Similarly, if we look at $[2t, 3t]$ and apply the mean value theorem, we get that there is some $c \in (2t, 3t)$ such that

$$\begin{aligned}
&\frac{f(3t) - f(2t)}{3t - 2t} = f'(c) \\
\Rightarrow &|f(3t) - f(2t)| = t|f'(c)| < t \cdot 4c < t \cdot 4 \cdot 3t = 12t^2 \\
\Rightarrow &|f(3t)| < 12t^2 + |f'(2t)| < 4t^2 + 8t^2 + 12t^2,
\end{aligned}$$

and (by a simple inductive argument) that if we look at $[(n-1)t, nt]$, we'll get the bound

$$|f(nt)| < 4t^2 + 8t^2 + 12t^2 + \ldots + 4nt^2 = 4(1 + 2 + \ldots + n)t^2$$

But we know that the sum of the first n positive integers is just $\frac{n(n+1)}{2}$: so we've shown that

$$|f(nt)| < 4 \cdot \frac{n(n+1)}{2} t^2 = 2t^2 \cdot (n)(n+1).$$

So: take any $y \in [0, 1]$, and write $y = n \cdot \frac{y}{n}$. Then, applying the above bound tells us that

$$|f(y)| = |f(n \cdot \frac{y}{n})| < 2 \left(\frac{y}{n}\right)^2 \cdot (n)(n+1) = 2y^2 \cdot \frac{n(n+1}{n^2} = 2y^2 \cdot \frac{n+1}{n}.$$

Letting n go to infinity tells us that $f(y) < 2y^2$. Identical calculations show that $p(1-y) > -2y^2 + 1$, and therefore that (in particular for $y = \frac{1}{2}$) we have $p(\frac{1}{2})$ both greater than and less than $\frac{1}{2}$, a contradiction. $\square$

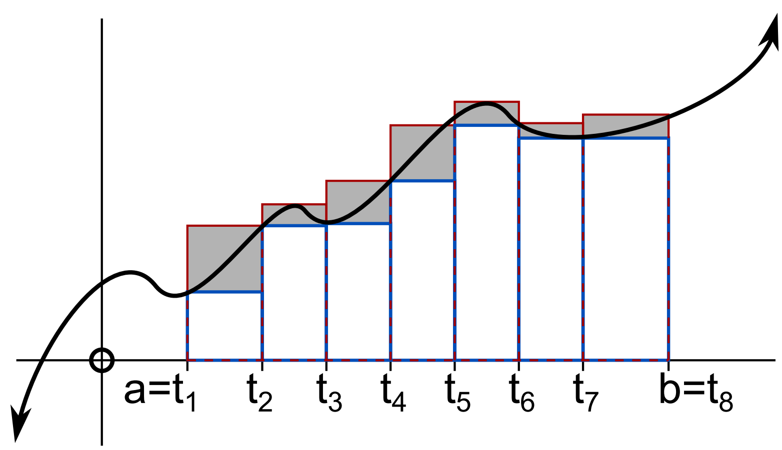
2 Integration

Definition. A function f is **integrable** on the interval $[a, b]$ if and only if the following holds:

- For any $\epsilon > 0$,
- there is a partition $a = t_1 < t_2 < \ldots < t_{n-1} < t_n = b$ of the interval $[a, b]$ such that

$$\left(\sum_{i=1}^{n} \sup_{x \in (t_i, t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i) - \sum_{i=1}^{n} \inf_{x \in (t_i, t_{i+1})} (f(x)) \cdot (t_{i+1} - t_i) \right) < \epsilon$$

One way to interpret the sums above is through the following picture:



Specifically,

- think of the ($\sum$inf)-sum as the area of the blue rectangles in the picture below, and
- think of the ($\sum$sup)-sum as the area of the red rectangles in the picture below.

- Then, the difference of these two sums can be thought of as the area of the gray-shaded rectangles in the picture above.
- Thus, we're saying that a function $f(x)$ is **integrable** iff we can find collections of red rectangles – an "upper limit" on the area under the curve of $f(x)$ – and collections of blue rectangles – a "lower limit" on the area under the curve of $f(x)$ – such that the area of these upper and lower approximations are arbitrarily close to each other.

Note that the above condition is equivalent to the following claim: if $f(x)$ is integrable, we can find a sequence of partitions $\{P_n\}$ such that "the area of the gray rectangles with respect to the P_n partitions goes to 0" – i.e. a sequence of partitions $\{P_n\}$ such that

$$\lim_{n\to\infty}\left(\sum_{i=1}^{n}\sup_{x\in(t_i,t_{i+1})}(f(x))\cdot(t_{i+1}-t_i)-\sum_{i=1}^{n}\inf_{x\in(t_i,t_{i+1})}(f(x))\cdot(t_{i+1}-t_i)\right)=0.$$

In other words, there's a series of partitions P_n such that these upper and lower sums both converge to the same value: i.e. a collection of partitions P_n such that

$$\lim_{n\to\infty}\sum_{i=1}^{n}\sup_{x\in(t_i,t_{i+1})}(f(x))\cdot(t_{i+1}-t_i)=\lim_{n\to\infty}\sum_{i=1}^{n}\inf_{x\in(t_i,t_{i+1})}(f(x))\cdot(t_{i+1}-t_i).$$

If this happens, then we define

$$\int_a^b f(x)dx=\lim_{n\to\infty}\sum_{i=1}^{n}\sup_{x\in(t_i,t_{i+1})}(f(x))\cdot(t_{i+1}-t_i)=\lim_{n\to\infty}\sum_{i=1}^{n}\inf_{x\in(t_i,t_{i+1})}(f(x))\cdot(t_{i+1}-t_i),$$

and say that this quantity is the **integral** of $f(x)$ on the interval $[a,b]$. For convenience's sake, denote the upper sums of $f(x)$ over a partition P as $U(f(x),P)$, and the lower sums as $L(f(x),P)$.

This discussion, hopefully, motivates why we often say that the integral of some function $f(x)$ is just "the area under the curve" of $f(x)$. Pictorially, we are saying that a function is integrable if and only if we can come up with a well-defined notion of area for this function; in other words, if sufficiently fine upper bounds for the area beneath the curve (the ($\sum$ sup)-sums) are arbitrarily close to sufficiently fine lower bounds for the area beneath the curve (the ($\sum$ inf)-sums.)

The definition of the integral, sadly, is a tricky one to work with: the sups and infs and sums over partitions amount to a ton of notation, and it's easy to get lost in the symbols and have no idea what you're actually manipulating. If you ever find yourself feeling confused in this way, just remember the picture above! Basically, there are three things to internalize about this definition:

- the area of the **red** rectangles corresponds to the upper-bound ($\sum$ sup)-sums,
- the area of the **blue** rectangles corresponds to the lower-bound ($\sum$ inf)-sums, and

- if these two sums can be made to be arbitrarily close to each other – i.e. the area of the gray rectangles can be made arbitrarily small – then we have a "good" idea of what the area under the curve is, and can say that $\int_a^b f(x)$ is just the limit of the area of those red rectangles under increasingly smaller partitions (which is also the limit of the area of the blue rectangles.)

The integral is a difficult thing to work with using just the definition: later on, we'll develop lots of tools to help us actually do nontrivial things with the integral. To illustrate how working with the definition goes, though, let's work two examples:

2.1 Calculating the Integral

Example. The integral of any constant function $f(x) = C$ from a to b exists; furthermore,

$$\int_a^b c dx = C \cdot (b - a).$$

Proof. Pick any constant function $f(x) = C$. To use our definition of the integral, we need to find a sequence of partitions P_n such that $\lim_{n\to\infty} U(f(x), P_n) - L(f(x), P_n)$ goes to 0. How can we do this? Well: what kinds of partitions of $[a, b]$ into n parts even exist?

One partition that often comes in handy is the uniform partition, where we break $[a, b]$ into n pieces all of the same length: i.e. the partition

$$P_n = \left\{a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n}, a + 3\frac{b-a}{n}, \ldots a + n\frac{b-a}{n} = b.\right\}$$

In almost any situation where you need a partition, this will work excellently! In particular, one advantage of this partition is that the lengths $(t_{i+1} - t_i)$ in the upper and lower sums are all the same: they're specifically $\frac{b-a}{n}$.

Let's see what this partition does for our integral. If we look at $U(f(x), P_n)$, where P_n is the uniform partition defined above, we have

$$\begin{aligned}
U(f(x), P_n) =& \sum_{i=0}^{n-1} \left(\sup_{x \in \left(a + i\frac{b-a}{n}, a + (i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(a + (i+1)\frac{b-a}{n} - a - i\frac{b-a}{n} \right) \\
=& \sum_{i=0}^{n-1} \left(\sup_{x \in \left(a + i\frac{b-a}{n}, a + (i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(\frac{b-a}{n} \right) \\
=& \sum_{i=0}^{n-1} C \cdot \left(\frac{b-a}{n} \right), \text{ because } f(x) \text{ is a constant function,} \\
=& C\frac{b-a}{n} + C\frac{b-a}{n} + \ldots + C\frac{b-a}{n} \\
&= C \cdot (b-a).
\end{aligned}$$

Similarly, if we look at $L(f(x), P_n)$, we have

$$\begin{aligned}
L(f(x), P_n) =& \sum_{i=0}^{n-1} \left(\inf_{x \in \left(a + i\frac{b-a}{n}, a+(i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(a + (i+1)\frac{b-a}{n} - a - i\frac{b-a}{n} \right) \\
=& \sum_{i=0}^{n-1} \left(\inf_{x \in \left(a + i\frac{b-a}{n}, a+(i+1)\frac{b-a}{n}\right)} f(x) \right) \cdot \left(\frac{b-a}{n} \right) \\
=& \sum_{i=0}^{n-1} C \cdot \left(\frac{b-a}{n} \right), \text{ because } f(x) \text{ is a constant function,} \\
=& C\frac{b-a}{n} + C\frac{b-a}{n} + \ldots + C\frac{b-a}{n} \\
&= C \cdot (b-a).
\end{aligned}$$

Therefore, we have that the limit

$$\lim_{n \to \infty} U(f(x), P_n) - L(f(x), P_n) = \lim_{n \to \infty} C \cdot (b-a) - C \cdot (b-a) = \lim_{n \to \infty} 0 = 0,$$

and consequently our integral exists and is equal to

$$\lim_{n \to \infty} U(f(x), P_n) = C \cdot (b-a).$$

□

Example. The function $f(x) = x^p$ is integrable on $[0, b]$ for any $p \in \mathbb{N}$ and $b \in \mathbb{R}^+$. Furthermore, the integral of this function is $\frac{b^{p+1}}{p+1}$.

Proof. Our uniform partition, where we broke our interval up into n equal parts, worked pretty well for us above! Let's see if it can help us in this problem as well. If we let $P_n = \{0, \frac{b}{n}, 2\frac{b}{n}, \ldots n\frac{b}{n} = b$, we have that $U(f(x), P_n)$ is just

$$\sum_{k=0}^{n-1} \sup_{x \in \left(k\frac{b}{n}, (k+1)\frac{b}{n}\right)} (x^p) \cdot \left((k+1)\frac{b}{n} - k\frac{b}{n} \right) = \sum_{k=0}^{n-1} \left((k+1)\frac{b}{n} \right)^p \cdot \frac{b}{n} = \frac{b^{p+1}}{n^{p+1}} \sum_{k=1}^{n} (k+1)^p,$$

and that the lower-bound sum, $(\sum \inf)$, is

$$\sum_{k=0}^{n-1} \inf_{x \in \left(k\frac{b}{n}, (k+1)\frac{b}{n}\right)} (x^p) \cdot \left((k+1)\frac{b}{n} - k\frac{b}{n} \right) = \sum_{k=0}^{n-1} \left(k\frac{b}{n} \right)^p \cdot \frac{b}{n} = \frac{b^{p+1}}{n^{p+1}} \sum_{k=1}^{n} k^p.$$

Taking the limit of their difference, we have that

$$\begin{aligned}
\lim_{n\to\infty} U(f(x),P_n) - L(f(x),P_n) &= \lim_{n\to\infty} \left(\left(\frac{b^{p+1}}{n^{p+1}} \sum_{k=0}^{n-1} (k+1)^p \right) - \left(\frac{b^{p+1}}{n^{p+1}} \sum_{k=0}^{n-1} k^p \right) \right) \\
&= \lim_{n\to\infty} \frac{b^{p+1}}{n^{p+1}} \left(\left(\sum_{k=0}^{n-1} (k+1)^p \right) - \left(\sum_{k=0}^{n-1} k^p \right) \right) \\
&= \lim_{n\to\infty} \frac{b^{p+1}}{n^{p+1}} \left((1^p + 2^p + 3^p + \ldots + n^p) - (0^p + 1^p + 2^p + \ldots + (n-1)^p) \right) \\
&= \lim_{n\to\infty} \frac{b^{p+1}}{n^{p+1}} \left(n^p \right) \\
&= \lim_{n\to\infty} \frac{b^{p+1}}{n} \\
&= 0.
\end{aligned}$$

Thus, by our definition, the function x^p is integrable on $[0,b]$! Furthermore, we have that the integral of this function is just

$$\lim_{n\to\infty} U(f(x),P_n) = \lim_{n\to\infty} \left(\frac{b^{p+1}}{n^{p+1}} \sum_{k=0}^{n-1} (k+1)^p \right) = \lim_{n\to\infty} \frac{b^{p+1}}{n^{p+1}} \left(\sum_{k=1}^{n} (k)^p \right).$$

So: it suffices to understand what the sum $(\sum_{k=1}^{n} (k)^p)$ is, for any p. Unfortunately, doing **that** is rather hard. We can find formulae for a few small cases: when $p=1$, for example, this is just

$$\lim_{n\to\infty} \frac{b^2}{n^2} \left(\sum_{k=1}^{n} k \right) = \lim_{n\to\infty} \frac{b^2}{n^2} \cdot \left(\frac{n(n+1)}{2} \right) = \lim_{n\to\infty} \frac{b^2}{2} \cdot \frac{n(n+1)}{n^2} = \frac{b^2}{2},$$

which indeed is the integral of x from 0 to b. In general, though, finding these sums is hard! Next week, we'll develop some tricks to get around this problem. $\square$