III. Integration

Economists seem much more interested in marginal effects and differentiation than in integration. Integration is important for finding the expected value and variance of random variables, which is used in econometrics as well as in individual decision making under uncertainty. There’s a lot of material to cover for integration, and anyhow, it’s an essential part of a calculus review.

Integration involves taking the sum of lots of infinitely tiny rectangles. Let’s say you have some function \( f(x) \) on some interval from \( a \) to \( b \), and you want to find the area between that function and the \( x \)-axis. If \( f \) is a “linear” function \( (f(x) = ax + b) \), this is no problem. If \( f \) is anything else, my inclination would be to break it up into something that looks like a staircase of sorts.

Let’s divide up the interval \([a,b]\) into \( n \) subintervals with the same length, so we have a series like \( a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b \), with \( x_i - x_{i-1} = \Delta x \). We use the value that \( f(x) \) takes at the right end of each interval to approximate its value on that whole interval. Then the area \( A \) under the curve is approximately the sum of all these little rectangles; each of which has width \( \Delta x \) and height \( f(x_i) \):

\[
A \approx \sum_{i=1}^{n} f(x_i) \Delta x
\]

As before with derivatives, a finer partitioning produces a better approximation. An infinitesimal partitioning gives a “perfect approximation.” When \( \Delta x \) is infinitesimal, we write \( dx \), and the summation sign is replaced with an integral:

\[
A = \int_{a}^{b} f(x) \, dx
\]

As a note, whatever comes between the \( \int \) and \( dx \) is sometimes called the integrand. Expected value problems sometimes use a discrete probability distribution, and sometimes a continuous distribution. Lifetime utility maximization problems add up the value of the utility function over a span of time; sometimes these area modeled using discrete time, and sometimes with continuous time. Each time, the discrete case involves a summation, and the continuous case an integral. For me, thinking of it this way makes these problems less intimidating.

The idea of integration is that almost every function is the derivative of some other function. If \( f \) is the derivative of \( F \), then \( F \) is called the antiderivative of \( f \). The function \( F \) is also what you get when you integrate \( f \):

\[
\int f(x) \, dx = F(x) + c \quad \text{is equivalent to:} \quad f(x) = F'(x)
\]

This is called an indefinite integral because no limits are specified, so we’re not evaluating it over any integral or space. We are left with an unidentified \( c \), called a constant of integration. When evaluating the integral over an interval, this constant goes away. Because it resolves the problem of this constant, an integral
over a predetermined interval is called a **definite integral**. This result is more or less the **first fundamental theorem of calculus**:

**Theorem:** Let $f$ be continuous on $[a,b]$. If $F$ is an antiderivative for $f$ on $[a,b]$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

That’s the theory. For practical matters, here are some of the basics rules of integration, to help you evaluate integrals (when they are solvable).

**Integrals are linear.** That is, constants can pass through the integral and there’s no problem with breaking the integral into additive parts:

$$\int [a \cdot f(x) + b \cdot g(x)]\,dx = a \int f(x)\,dx + b \int g(x)\,dx$$

The **power rule in reverse works most of the time.** That is:

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + c,$$

provided that $n \neq -1$. (If $n = 1$, recall the natural logarithm.)

**Example:**

$$\int (14x^2 + 7\sqrt{x} - 9)\,dx$$

**When you have a function next to its own derivative, try a change of variables.** If you are integrating some complicated function of $x$ with respect to $x$, but you notice that it looks as if the integrand contains something multiplied by its derivative, this formula might be helpful:

$$\int f(u)\,du = \int [f(g(x)) \cdot g'(x)]\,dx$$

This is the opposite of the chain rule.

**Example:**

$$\int \frac{x}{\sqrt{3x^2 + 6}}\,dx$$

**Integration by parts is your friend.** You might have to integrate some complicated function, but you notice that one part looks like the derivative of an easy function, whereas the other part looks like it has a pretty straightforward derivative.

$$\int [f(x) \cdot g'(x)]\,dx = f(x) \cdot g(x) - \int [f'(x) \cdot g(x)]\,dx$$

This is essentially a double change of variable.

**Example:**

$$\int xe^x\,dx$$

Additionally, you should keep in mind all the “special derivatives,” like those of $f(x) = \ln x$ and $f(x) = e^x$, when integrating. On top of these, there are many
functions with special antiderivatives that you have to memorize, be able to derive, or at least recognize. Trigonometric functions are some of these. Functions with lots of squares and square roots might be as well. Chapter 9 of Sydsæter, Strøm, and Berck has lists many of the functions and their antiderivatives. Most basic calculus textbooks also have a table of integrals. Since I’ve never encountered these in a couple of years of grad school, I’m not going to cover them.

There are a few things to remember about definite integrals. First of all, switching the order of the limits changes the sign of the integral. Second, the integral over a single point has a value of zero (this comes up in probability). Third, it’s fine to split up the limits of integration:

$$\int_a^b f(x) \, dx = -\int_b^a f(x) \, dx$$
$$\int_a^a f(x) \, dx = 0$$
$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx$$

Some other useful properties relate to taking derivatives of integrals. The first two say what happens when you differentiate with respect to the limits of integration. The next is for differentiating with respect to a parameter inside the integral. The last is Leibnitz’ rule, a generalization of all these, that also tells you what to do when differentiating with respect to the independent variable of \( f \):

$$\frac{d}{db} \int_a^b f(x) \, dx = f(b)$$
$$\frac{d}{da} \int_a^b f(x) \, dx = -f(a)$$

$$\frac{d}{dt} \int_a^b f(x,t) \, dx = \int_a^b \frac{\partial f(x,t)}{\partial t} \, dx \quad \text{(when } a, b, \text{ and } x \text{ do not depend on } t\text{)}$$

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x,t) \, dx = f(v(t),t)v'(t) - f(u(t),t)u'(t) + \int_{a(t)}^{b(t)} \frac{\partial f(x,t)}{\partial t} \, dx$$

The integrals we have talked about have been for functions defined on closed intervals, which are compact sets. An integral over an interval from \( a \) to \( b \) in which either \( a \) or \( b \) (or both!) equals \( \pm\infty \) or in which the integrand \( f \) is undefined at some point in \([a,b]\) is an improper integral.

**Example:**  \( \int_1^\infty \frac{dx}{x^2} \) because the upper limit is infinity

**Example:**  \( \int_{-1}^1 \frac{xdx}{x} \) because the function \( \frac{x}{x} \) isn’t defined at \( x = 0 \)
In these cases, the trick is to stick in some constant like \( c \) for the offending value, and then take evaluate the limit of the integral as \( c \) approaches that value. If the value is in the middle of the integral, you have to split the integral into multiple parts.

Example:
\[
\lim_{c \to \infty} \int_1^c \frac{dx}{x^2} = \lim_{c \to \infty} \left[ \frac{-1}{x} \right]_1^c = \lim_{c \to \infty} \left( 1 - \frac{1}{c} \right) = 1
\]

Example:
\[
\lim_{c \to 1} \int_1^c \frac{xdx}{x} = \lim_{c \to 1} \int_1^c \frac{xdx}{x} + \lim_{c \to 0} \int_0^c \frac{x dx}{x}
\]

Integration will frequently be used to find the expected value of a function, often expected utility. Though probability is going to be covered on the last day of this course, we’ll do expectations briefly now, so we can work some economic examples.

Example: My friend Oliver and I have a bet. He flips a coin and if it comes up heads, he wins two dollars from me. If it comes up tails, I win one dollar from him. What is my expected value of this bet? Calculating that I’d get $1.50 50% of the time, plus negative $3.50 50% of the time, I’d guess that this is a losing prospect in general. And I’d be correct, but I’d like to know how to state this formally.

Let’s let \( X \) be some random variable, something whose value is determined by chance. The set \( S \) will denote all possible values of the outcome of this random variable. (In this example, \( X \) can be the amount that I win, and the set of all possible outcomes is \( S = \{-3, +1\} \).) There is some probability distribution \( P \) that determines the likelihood of each outcome (\( P[X = -3] = 0.5 \) and \( P[X = +1] = 0.5 \)). The expected value of \( X \) is defined as:

\[
\mathbb{E}[X] = \sum_{s \in S} s P[X = s]
\]

That is, the sum over all possible values of the outcome, times the chance of that outcome occurring. This will give you the average value of \( X \) over many, many independent repetitions of the experiment.

We can also take the expected value of some function of \( X \). For example, I might get some utility from my monetary winnings, represented by \( U(X) \); I want to know my expected utility. This would be calculated as:

\[
\mathbb{E}[U(X)] = \sum_{s \in S} U(s) P[X = s]
\]

There is nothing special about the function \( U \); any \( g(X) \) or \( h(X) \) will do.

Example: My utility function is \( U(c) \), where \( c \) is my consumption, equal to my initial wealth plus or minus any winnings. I start with $10, and with equal probability, I either lose $3 or win $1. My expected winnings are:
\[ E[X] = (0.5)(-3) + (0.5)(1) = -1 \]

*My expected consumption is:*

\[ E[10 + X] = (0.5)(10 - 3) + (0.5)(10 + 1) = 9 \]

*and my expected utility is:*

\[ E[\ln(10 + X)] = (0.5)\ln(10 - 3) + (0.5)\ln(10 + 1) = 0.5(\ln 7 + \ln 9) \]

When we have a continuous set of outcomes (say, we flip the coin an trillion times and look at the percentage of heads and tails), then there is a **probability density function** \( f \) with the property that for any subset \( T \) of \( S \),

\[ P(X \in T) = \int_T f(s)\,ds \]

Notice that with continuous distributions, the probability that you get any one element \( s \) of set \( S \) is zero, since the integral from \( s \) to \( s \) is zero. On the other hand, the probability that you get *something* in \( S \) must be exactly one; that is, \( \int_S f(s)\,ds = 1 \) (this is a property that permissible PDFs must have). Analogous to the discrete case, the expected value of a random variable is:

\[ E[X] = \int_{s \in S} sf(s)\,ds \]

and the expected value of some function \( h(X) \) is defined as:

\[ E[h(X)] = \int_{s \in S} h(s)f(s)\,ds \]

**Example:** The number of years it takes grad students to complete a PhD is distributed \( f(t) = 0.187e^{-1.87t} \), where \( t \geq 0 \). How long can an incoming grad student expect to be in grad school?

**Example:** Your utility function for the number of years spent in grad school is:

\[ U(t) = C - t^2 \]

(You might graduate in four years with little disutility; it might take you a few decades, in which case it is very painful.) Before you start, you want to know: what is the expected utility of grad school?

**Example:** The starting salaries of new PhDs in economics is distributed:

\[ f(y) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) \]
If $\mu = 56,412$ and $\sigma = 8,273$, then what is their expected starting salary?

**Example:** The utility a person gets from this salary is determined by the function:

$$U(c) = \alpha c - \beta c^2 + \gamma$$

What's his expected utility?

**References:**

Simon and Blume: Appendix A.4.

Sydsæter, Strøm, and Berck: Chapter 9.

Salas and Hille: Chapters 5, 8, and 17.