Fall Semester '07-'08
Akila Weerapana

LECTURE 5: INTEGRAL CALCULUS

I. INTRODUCTION

- In the last lecture, we talked about differentials and used differentials to talk about some important economic concepts like, elasticities, economic growth and simple comparative static analysis. Since Econ 205 is a pre-requisite, that will be all we will cover on differential calculus theory. You should be very comfortable with any other topic related to differential calculus at the level of Math 205.
- Today's lecture does a similar review of concepts regarding integral calculus from Math 205. Once again, we will do a quick review and spend most of the time discussing economic applications. The most important applications will be calculating consumer and producer surplus, bond pricing and some econometric applications with density functions.
- If you need to brush up on your integrals, once again I urge you to do so immediately. Chapter 12 of Klein will serve both as a review of what you studied in your calculus classes and as supplementary reading for this lecture.

II. THEORY

- The **indefinite integral** $F(x) = \int f(x)dx$ is the (family of) anti-derivative(s) of a function. In other words, $\frac{dF(x)}{dx} = f(x)$.
- Since constant terms have a derivative of zero, we can only define $F(x)$ up to an arbitrary constant absent any further information, hence the family of anti-derivatives.
- In essence, you can think of **integration** as the reverse of differentiation. Suppose that we have a function $f(x)$, that is the derivative of some function $F(x)$. Integration can be used (with a little bit of additional information to uncover the constant terms) to recover the original function $F(x)$ using the information contained in $f(x)$.
- The **definite integral** $\int_a^b f(x)dx$ is the area under the curve f(x) over the range $x = a$ to $x = b$.
- The value of $\int_a^b f(x)dx = F(b) - F(a)$ where $F(x) = \int f(x)dx$
- In Economics, this means that we can derive a total cost function using information about marginal costs, derive the underlying utility function given information about marginal utility and derive a cumulative density function given information about a probability density function, etc.

Basic Rules of Indefinite Integrals

- Since almost everyone uses integration less than differentiation in academic settings, it may be wise to do a quick survey of some basic rules of integration. Some useful rules to remember (where a is an arbitrary constant term) are

$$\int x^n dx = \frac{x^{n+1}}{n+1} + c \qquad \int a dx = ax + c$$
$$\int e^x dx = e^x + c \qquad \int \frac{1}{x} dx = \ln(|x|) + c$$
$$\int [f(x) + g(x)] dx = \int f(x)dx + \int g(x)dx$$
$$\int af(x)dx = a \int f(x)dx$$

- Note that c is an arbitrary constant. Since constants disappear during differentiation without additional information we can not identify the constant terms in the original function by integration, i.e. given $f'(x) = 2x$ we can not state whether $f(x) = x^2 + 3$ or $f(x) = x^2 + 5$

- There are also two important rules of substitution that you should be familiar with.

  1. If u=g(x) then $\int f(u)g'(x)dx = \int f(u)du$
  2. $\int udv = uv - v \int du$

  Examples:

  1. Find the value of the following integral $\int \frac{2x}{x^2+3} dx$. Let $u = g(x) = x^2 + 3$. Then $du = g'(x)dx = 2xdx$. We can then rewrite the integral as $\int \frac{1}{u} du$ the solution to which is $\ln(|u|) + c$. So
  $$\int \frac{2x}{x^2+3} dx = \ln(x^2 + 3) + c$$
  2. Find the value of the following integral $\int \ln(x)dx$. Let $u = \ln(x)$ and $v = x$. Then $du = (1/x)dx$ and $dv = dx$. Since we have an integral of the form $\int udv$ we have a solution of the form $uv - \int vdu = x\ln(x) - \int x \left(\frac{1}{x}dx\right) = xln(x) - \int dx$ so
  $$\int \ln(x)dx = x\ln(x) - x + c$$

Basic Rules of Definite Integrals

- Other important rules to keep in mind are

$$\int_a^a f(x)dx = 0$$
$$\int_a^b f(x)dx = -\int_b^a f(x)dx$$
$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx \text{ where } a < b < c$$

III. APPLICATIONS

Recovering Total Cost from Marginal Cost

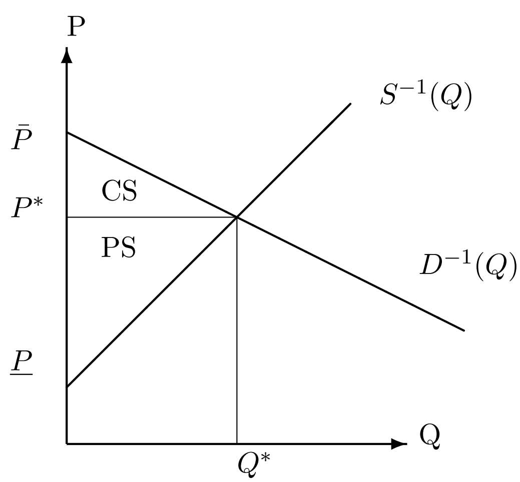
- Given a marginal cost function of the form MC(Q) we can recover the underlying total cost function as $TC(Q) = \int MC(Q)dQ$. For example, if we have a marginal cost function of the form $MC(Q) = 3 + 2Q$, the underlying total cost function is $TC(Q) = 3Q + Q^2 + c$.

- Note that the total cost function can only be pinned down upto a constant. What is the constant in a total cost function? It is fixed costs, so using a marginal cost function we are unable to recover the level of fixed costs that a firm faces.

Consumer and Producer Surplus

- In microeconomics, integrals play a vital role in calculating consumer and producer surplus. This is especially true when the demand curve and supply curve are not linear. Let's think about how to determine consumer and producer surplus using the supply/demand diagram below.

- Let the demand function be denoted by $Q = D(P)$ and the supply function be denoted by $Q = S(P)$. The equilibrium price and quantity, i.e. where demand and supply intersect are denoted as P^* and Q^* respectively.



- Consumer surplus is the area under the demand curve above P^*, i.e. the difference between willingness to pay and price paid for each unit sold up to Q^*. This area is labeled CS in the above diagram.

- Producer surplus is the area under the supply curve below P^*, i.e. the difference between the marginal cost of production (represented by the supply curve) and the market price, which represents the producer's profits. This area is labeled PS in the above diagram.

- We can write down algebraic expressions for producer and consumer surplus using integrals. For example consumer surplus is the area under the demand curve and above P^* between $Q = 0$ and $Q = Q^*$. We can write this in one of two ways:

$$
\begin{aligned}
CS &= \int_0^{Q^*} \left[D^{-1}(Q) - P^*\right] dQ \\
\text{or equivalently } CS &= \int_{P^*}^{\bar{P}} \left[D(P)\right] dP
\end{aligned}
$$

- Producer surplus is the area between P^* and the supply curve between $Q = 0$ and $Q = Q^*$.

$$
\begin{aligned}
PS &= \int_0^{Q^*} \left[P^* - S^{-1}(Q)\right] dQ \\
\text{or equivalently } PS &= \int_{\underline{P}}^{P^*} \left[S(P)\right] dP
\end{aligned}
$$

Example:

- Suppose the demand curve for a product is given by $D(P) = 10 - 2P$ and the supply curve is given by $S(P) = 4 + P$. Equilibrium price and quantity can be calculated as $P^* = 2$ and $Q^* = 6$. The inverse demand and supply functions can be calculated as $D^{-1}(Q) = 5 - \frac{Q}{2}$ and $S^{-1}(Q) = Q - 4$
- The y-intercepts are at $\bar{P} = 5$ for demand and $\underline{P} = -4$ for supply.
- Using integrals, we can calculate consumer surplus to be

$$
\begin{aligned}
CS &= \int_0^{6} \left[D^{-1}(Q) - 2\right] dQ \\
&= \int_0^{6} \left[\left(5 - \frac{Q}{2}\right) - 2\right] dQ = \int_0^{6} \left[3 - \frac{Q}{2}\right] dQ \\
&= \left(3Q - \frac{Q^2}{4}\right)\Bigg|_0^6 = 9
\end{aligned}
$$

- Alternatively, we could have found consumer surplus as

$$
\begin{aligned}
CS &= \int_2^{5} \left[D(P)\right] dP \\
&= \int_2^{5} \left[10 - 2P\right] dP \\
&= \left(10P - P^2\right)\Big|_2^5 = (50 - 25) - (20 - 4) = 9
\end{aligned}
$$

- Using integrals, we can calculate producer surplus to be

$$\begin{aligned} PS &= \int_0^6 \left[2 - S^{-1}(Q)\right] dQ \\ &= \int_0^6 \left[2 - (Q-4)\right] dQ = \int_0^6 [6 - Q]\, dQ \\ &= \left(6Q - \frac{Q^2}{2}\right)\Bigg|_0^6 = 18 \end{aligned}$$

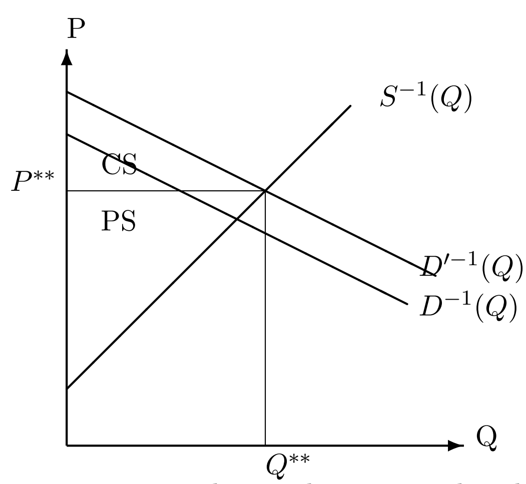
- Alternatively, we could have found producer surplus as

$$\begin{aligned} PS &= \int_{-4}^2 [S(P)]\, dP \\ &= \int_{-4}^2 [4 + P]\, dP \\ &= \left(4P + \frac{P^2}{2}\right)\Bigg|_{-4}^2 = (8+2) - (-16+8) = 18 \end{aligned}$$

- Of course, since the curves are linear, we can avoid all of this and calculate the magnitudes of consumer and producer surplus geometrically. Consumer surplus is the area covered by a right-angled triangle with base 6 and height 3 so $CS = 1/2 * 6 * 3 = 9$. Producer surplus is the area under a right-angled triangle with base 6 and height 6 so $PS = 1/2 * 6 * 6 = 18$.

Welfare Effects of Price Changes

- We can also use integrals to think about the welfare effects of changes in equilibrium price and quantity. For example: what happens to consumer and producer surplus when price goes up? Graphically, let's think about what happens when there is an increase in demand that raises equilibrium price to P^{**} and equilibrium quantity to Q^{**}.



- New consumer surplus is the area under the new demand curve above P^{**} and the new producer surplus is the area under the supply curve below P^{**}.

  Example:

- Suppose the new demand curve (D') for the product is $Q = 13 - 2P$ and the supply curve is still given by $Q = 4 + P$. New equilibrium price and quantity can be calculated as $P^{**} = 3$ and $Q^{**}= 7$.
- Consumer surplus is now $CS = \int_0^7 \left[D'^{-1}(Q) - 3\right] dQ$. The new demand curve can be inverted as $P = (13 - Q)/2$, so $CS = \int_0^7 \left[\frac{(13-Q)}{2} - 3\right] dQ = \int_0^7 \left[\frac{(7-Q)}{2}\right] dQ$ which simplifies to

$$CS = \left(\frac{7Q}{2} - \frac{Q^2}{4}\right)\Big|_0^7 = \frac{49}{4}$$

- Producer surplus is now $PS = \int_0^7 \left[3 - S^{-1}(Q)\right] dQ = \int_0^7 \left[3 - (Q - 4)\right] dQ$ which simplifies to

$$PS = \int_0^7 (7 - Q)dQ = \left(7Q - \frac{Q^2}{2}\right)\Big|_0^7 = \frac{49}{2}$$

- Consumer surplus increases from 9 to 12.25 and producer surplus increases from 18 to 24.5.

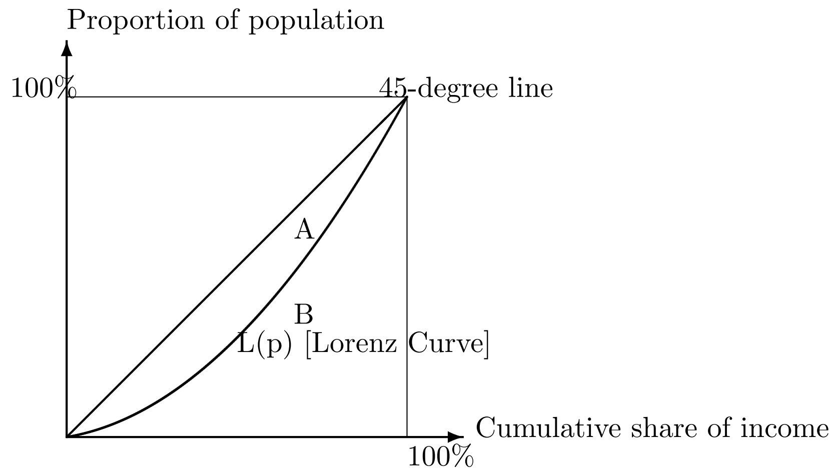
Gini Coefficients

- Another useful application of integrals described in Klein is the calculation of a **Gini Coefficient**, a widely used measure of income inequality.
- The Gini coefficient is derived using integral calculus from a graph known as a Lorenz curve. A Lorenz curve simply plots the cumulative share of overall income earned by each percentile of the population against each percentile of the population, ordered from poorest to richest.
- A 45-degree line is drawn on the diagram to represent complete equality - the case where every percentile of the population earns the same share of overall income. The more equal the income distribution is, the closer the Lorenz curve is to matching the 45-degree line.
- The Gini coefficient is defined as the area between the 45-degree line and the Lorenz curve over the area under the 45 degree line. In terms of the graph below we can define

$$\text{Gini} = \frac{A}{A + B}$$

- Mathematically, the Gini coefficient can be calculated as

$$\begin{aligned} G &= \frac{\int_0^1 [p - L(p)]\, dp}{\int_0^1 p dp} \\ &= \frac{\frac{p^2}{2}\Big|_0^1 - \int_0^1 [L(p)]\, dp}{\frac{p^2}{2}\Big|_0^1} = \frac{\frac{1}{2} - \int_0^1 [L(p)]\, dp}{\frac{1}{2}} \\ &= 1 - 2\int_0^1 L(p) dp \end{aligned}$$



Random Variables

- We can also use integrals to think about basic concepts related to econometrics and statistics. In statistics, we often use random variables, variables whose value reflect the outcome of some probabilistic event.

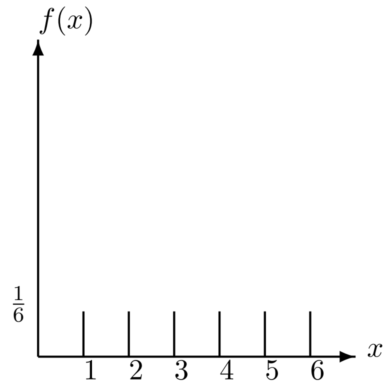
- Random variables have a probability distribution, which describes the values that the random variable can take on, and the probability of achieving each of those outcomes. For example, suppose X is a random variable whose value, x, is the outcome from a single roll of a 6 sided fair die. The probability distribution of X, f(x) is f(1)=1/6, f(2)=1/6, f(3)=1/6, f(4)=1/6, f(5)=1/6, f(6)=1/6 and f(i)=0 for any other value i.

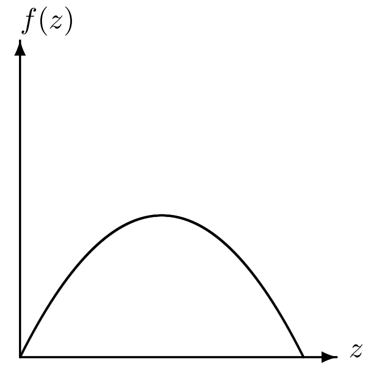
- For any discrete random variable X with the probability distribution function f(x), the following rules must hold $0 \leq f(x) \leq 1$ and $\sum_{\forall x_i} f(x_i) = 1$ i.e. the probability of observing any outcome has to be non-negative and the sum of the probability of observing the independent outcomes can't exceed 1.

- Random numbers can be continuous as well as discrete. Suppose that Z is a continuous random variable that can take on a continuum of values z. Since z can take on an infinite number of values, we can't describe the probability of taking any given value: we can only talk about the probability that z taking on a range of values as $Pr(a \leq z \leq b) = \int_a^b f(z)dz \geq 0$

- As in the discrete case, the second condition states that all the probabilities have to add up to 1: so $\int_{-\infty}^{\infty} f(z)dz = 1$. Note that $f(i) = 0$ if Z does not take on the value i.

- Possible distribution functions for a discrete random variable X and a continuous random variable Z, are given below.





- Other key concepts associated with random variables include the cumulative distribution function $F(x)$ defined as

$$F(x) = Prob[X \leq x] = \int_{-\infty}^{x} f(x)dx$$

- The expected value of a random variables defined as

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

- The variance of a random variable defined as

$$E(x-\mu)^2 = E(x^2 - 2\mu x + \mu^2) = E(x^2) - 2\mu(E(x)) + \mu^2 = E(x^2) - \mu^2$$

- This can be calculated as

$$Var(x) = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

Examples

- Let's do some examples with a couple of commonly used probability distributions highlighted by Klein: the uniform distribution and the exponential distribution.

- The uniform distribution has an upper limit (b) and a lower limit (a) and has a distribution function

$$\begin{aligned} f(x) &= \frac{1}{b-a} \text{ for } x \text{ in } [a,b] \\ f(x) &= 0 \text{ otherwise} \end{aligned}$$

- The CDF for the uniform distribution is then

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(x)dx = \int_{a}^{x} \left[\frac{1}{b-a}\right] dx \\ &= \left.\frac{x}{b-a}\right|_{a}^{x} \\ F(x) &= \frac{x-a}{b-a} \end{aligned}$$

- The expected value of a random variable that is uniformly distributed is

$$\begin{aligned} E(x) &= \int_{-\infty}^{\infty} xf(x)dx = \int_{a}^{b} \left[\frac{x}{b-a}\right] dx \\ &= \left(\frac{1}{b-a}\right) \left.\frac{x^2}{2}\right|_{a}^{b} \\ &= \frac{b^2-a^2}{2(b-a)} \\ E(x) &= \frac{b+a}{2} \end{aligned}$$

- The variance of a random variable that is uniformly distributed is

$$\begin{aligned} Var(x) &= E(x-\mu)^2 = E(x^2) - \mu^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - \frac{(b+a)^2}{4} = \int_{a}^{b} x^2 f(x)dx - \frac{(b+a)^2}{4} \\ Var(x) &= \frac{1}{b-a}\int_{a}^{b} x^2 dx - \frac{(b+a)^2}{4} = \frac{1}{b-a}\left.\frac{x^3}{3}\right|_{a}^{b} - \frac{(b+a)^2}{4} \\ &= \frac{(b^3-a^3)}{3(b-a)} - \frac{(b+a)^2}{4} = \frac{(b^2+ab+a^2)}{3} - \frac{(b^2+2ab+a^2)}{4} \\ &= \frac{(b^2-2ab+a^2)}{12} \\ Var(x) &= \frac{(b-a)^2}{12} \end{aligned}$$

- A random variable that is exponentially distributed has a distribution function

$$f(x) = \lambda e^{-\lambda x} \text{ for } 0 < x < \infty$$

- The CDF for the exponential distribution is then

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(x)dx = \int_{0}^{x} \left[\lambda e^{-\lambda x}\right] dx \\ &= \left.-e^{-\lambda x}\right|_{0}^{x} \\ F(x) &= 1 - e^{-\lambda x} \end{aligned}$$

- The expected value of a random variable that is exponentially distributed is

$$
\begin{aligned}
E(x) &= \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{\infty} \left[x\lambda e^{-\lambda x}\right] dx \\
& \quad \text{Define } u = x, dv = \lambda e^{-\lambda x} dx \text{ then} \\
& \quad du = dx, v = -e^{-\lambda x} \\
E(x) &= uv - \int vdu = -xe^{-\lambda x}\Big|_0^{\infty} - \int_0^{\infty} \left[-e^{-\lambda x}\right] dx \\
&= -xe^{-\lambda x}\Big|_0^{\infty} - \frac{e^{-\lambda x}}{\lambda}\Big|_0^{\infty} \\
&= -\left(x + \frac{1}{\lambda}\right) e^{-\lambda x}\Big|_0^{\infty} \\
&= -\left(x + \frac{1}{\lambda}\right) e^{-\lambda x}\Big|_0^{\infty} \\
&= \frac{1}{\lambda}
\end{aligned}
$$

- The variance of a random variable that is exponentially distributed is

$$
\begin{aligned}
Var(x) &= E(x-\mu)^2 = E(x^2) - \mu^2 \\
&= \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2
\end{aligned}
$$

We can calculate the value of the first term in the above expression as

$$
\begin{aligned}
& \quad \text{Define } u = x^2, dv = \lambda e^{-\lambda x} dx \text{ then} \\
& \quad du = 2xdx, v = -e^{-\lambda x} \\
\int_0^{\infty} \left[x^2 \lambda e^{-\lambda x}\right] dx &= uv - \int vdu = -x^2 e^{-\lambda x}\Big|_0^{\infty} - \int_0^{\infty} \left[-e^{-\lambda x}\right] (2x)dx \\
&= -x^2 e^{-\lambda x}\Big|_0^{\infty} + \frac{2}{\lambda}\int_0^{\infty} \left[\lambda x e^{-\lambda x}\right] dx \\
&= -x^2 e^{-\lambda x}\Big|_0^{\infty} + \left(\frac{2}{\lambda}\right)\left(\frac{1}{\lambda}\right) \\
\int_0^{\infty} \left[x^2 \lambda e^{-\lambda x}\right] dx &= \left(\frac{2}{\lambda^2}\right)
\end{aligned}
$$

Therefore

$$
Var(x) = \left(\frac{2}{\lambda^2}\right) - \left(\frac{1}{\lambda^2}\right) = \frac{1}{\lambda^2}
$$