Section Notes 4

Wonbin Kang

October 1, 2009

Agenda

1. Linear Approximation of EV (or CV)
2. Probability Review[1]
3. Expected Utility
4. Independence Axiom

1 Linear Approximation of EV

Problem 1. Assume that there are two goods in the world. For a general change in the price of good 1 (while the price of good 2 remains the same), calculate the EV (or CV). Assume that the Walrasian demand function for good 1 given a particular price-wealth vector is equal to $x_1(p_1, p_2, w)$, and the Hicksian demand function good 1 given a particular price-utility vector is equal to $h_1(p_1, p_2, u)$. For purposes of this exercise, assume that the Walrasian and Hicksian demand functions are *not* linear and further assume that you can only observe the price-wealth vectors, but not the utilities of the agents.

We're faced with two problems: first, we have a non-linear function which makes analytical problem solving very difficult, and two, if we wanted to empirically test our model, solving for EV requires us to use a parameter that is not observable, u^1, or u^0 if we want to calculate CV.

1.1 Taylor Approximation

A basic idea in mathematics when dealing with non-linear functions is to approximate such an equation with a linear function. A tool for such approximation is the first-order Taylor approximation. Consider a non-linear function $f : \Re^n \to \Re$. We can approximate this function around the point $\vec{x}^*$:

[1] I'm going to assume that most of you know basic probability. If you're having trouble, please schedule office hours with me or Sam.

$$f(\vec{x}) \approx \tilde{f}(\vec{x}) = f(\vec{x}^*) + \nabla f(\vec{x}^*)(\vec{x} - \vec{x}^*)$$

, or, if we're dealing with a function $f : \Re \to \Re$ (as we will be dealing with when approximating EV and CV)

$$f(x) \approx \tilde{f}(x) = f(x^*) + f'(x^*)(x - x^*)$$

. This gives us a linear approximation of the function f, using a linear function that is tangent to f at the point $\vec{x}^*$ or x^*. What is important to see is that the approximation becomes worse as the values of $\vec{x}$ move away from $\vec{x}^*$. Therefore, it is important that you choose the "optimal" $\vec{x}^*$.

1.2 Application to Estimation of EV

Now recall that, we want to calculate EV, which is equal to $\int_{p_1^1}^{p_1^0} h_1(p_1, \vec{p}_{-1}, u^1)dp_1 = \int_{p_1^1}^{p_1^0} h_1(p_1, p_2^0, u^1)dp_1 (\because \vec{p} \in \Re^2)$. So if we know that the Hicksian demand function of good 1 is non-linear, then we approximate it using the first-order Taylor Approximation. First, what value of $p_1$should we use to approximate the non-linear Hicksian? The answer is obviously the new price of good 1, or p_1^1, because we are using the Hicksian demand curve for the new utility level u^1, or the Hicksian demand curve through the point $p_1^1$and $h_1(p_1^1, p_2^0, u^1)$.[2] And so:

$$h_1(p_1, p_2^0, u^1) \approx h_1(p_1^1, p_2^0, u^1) + \frac{\partial h_1(p_1^1, p_2^0, u^1)}{\partial p_1}(p_1 - p_1^1). \tag{1}$$

The problem is that equation 1 cannot be observed. Can you observe u^1? Therefore, we need to change the approximation a function that includes only parameters that we can observe. First, we know that:

$$h_1(p_1^1, p_2^0, u^1) = x_1(p_1^1, p_2^0, w) \tag{2}$$

. [3]Second, we can use the Slutsky equation for the partial derivative in the second argument of equation 1:

$$\frac{\partial h_1(p_1^1, p_2^0, u^1)}{\partial p_1} = \frac{\partial x_1(p_1^1, p_2^0, w)}{\partial p_1} + \frac{\partial x_1(p_1^1, p_2^0, w)}{\partial w} \times x_1(p_1^1, p_2^0, w). \tag{3}$$

If we plug in equations **??** and 3 into equation 1, then we have the linear estimation of the non-linear Hicksian demand curve for good 1, in terms of observable parameters[4]:

[2]So what point should we use if we were estimating the CV of a price change in good 1? Do you see that the approximation has to be around the point $(p_1^0, h_1(p_1^0, p_2^0, u^0))$?

[3]If you don't see this, try drawing a picture.

[4]In other words, we've gotten rid of the pesky u^1.

$$h_1(p_1, p_2^0, u^1) \approx x_1(p_1^1, p_2^0, w) + \left[\frac{\partial x_1(p_1^1, p_2^0, w)}{\partial p_1} + \frac{\partial x_1(p_1^1, p_2^0, w)}{\partial w} \times x_1(p_1^1, p_2^0, w) \right] (p_1 - p_1^1). \tag{4}$$

One final note, notice the difference between the estimation of the Hicksian demand curve in equation 4 and the following Taylor approximation of the Walrasian demand curve around the point $(p_1^{1‘}, x_1(p_1^1, p_2^0, w))$ (which we could use to calculate the Marshallian Consumer Surplus):

$$x_1(p_1, p_2^0, u^1) \approx x_1(p_1^1, p_2^0, u^1) + \frac{\partial x_1(p_1^1, p_2^0, u^1)}{\partial p_1}(p_1 - p_1^1). \tag{5}$$

Try this for CV.

1.3 Example (from Miller Notes p.81)[5]

Suppose that we knew that $x_1(p_1^0, p_2^0, w) = 100$; $p_1^0 = 10$; $\frac{\partial x}{\partial p} = -4$; and $\frac{\partial x}{\partial w} = 0.02$. Let's assume that the price of good 1 increased to $p_1^1 = 12.5$. The question asks: "How much should a public assistance program aimed at maintaining a certain standard of living be increased to offset this price increase?"

Obviously, we need to calculate the CV (since the question was phrased in terms of a social planner compensating the consumers for a price increase[6]).

Since we're looking for the CV, we have to use the Hicksian demand correspondence for the utility level u^0, $h_1(p_1, p_2^0, u^0)$. We know that $h_1(p_1^0 = 10, p_2^0, u^0) = x_1(p_1^0 = 10, p_2^0, w) = 100$, and we know that we can approximate the slope of $h_1(p_1, p_2^0, u^0)$ via the Slutsky equation as follows:

$$\begin{aligned}
\frac{\partial h_1(p_1^0, p_2^0, u^0)}{\partial p_1} &= \frac{\partial x_1(p_1^0, p_2^0, w)}{\partial p_1} + \frac{\partial x_1(p_1^0, p_2^0, w)}{\partial w} \times x_1(p_1^0, p_2^0, w) \\
&= -4 + 0.02 \times 100 \\
&= -2
\end{aligned}$$

. Therefore, the Hicksian demand at the new price $p_1^1 = 12.5$ is approximated by:

$$\begin{aligned}
h_1(p_1, p_2^0, u^0) &\approx h_1(p_1^0, p_2^0, u^0) + \frac{\partial h_1(p_1^0, p_2^0, u^0)}{\partial p_1}(p_1 - p_1^0) \\
h_1(p_1, p_2^0, u^0) &\approx 100 + (-2) \times (p_1 - 10) \\
\therefore h_1(p_1 = 12.5, p_2^0, u^0) &\approx 100 + (-2) \times 2.5 = 95
\end{aligned}$$

. Computing the CV, we get:

[5] Since much of the explanation can be found in the Miller Notes, my recitation below will be brief.

[6] You should know that CV is going to be negative.

$$|CV| = (2.5) \times \left(\frac{95+100}{2}\right) = 243.75.$$

As an exercise, try calculating the Marshallian Consumer Surplus (ΔCS) by use of the linear approximation of the Walrasian demand function.

2 Probability Review

2.1 PDF (PMF): Probability Density (Mass) Function

In the discrete case, the PMF of a random variable X is given by $f(x) = \Pr(X = x)$. Note that the PMF $f(x)$ must be defined from all values of x that can be taken by the random variable X.

Example 2. $f(x) = \begin{cases} \frac{1}{2}\, if & X = 5 \\ \frac{1}{2}\, if & X = 10 \\ 0 & otherwise \end{cases}$

In the continuous case, the PDF of a random variable X is also given by $f(x)$, however, in this case, the $\Pr(X = x) = 0$. The PDF is defined in terms of intervals and so, $\Pr(a \leq X \leq b) = \int_a^b f(x)dx$. Note that $\int_{-\infty}^{\infty} f(x)dx = 1$.

Example 3. $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ is the pdf of a normal distribution with mean $\mu$and variance equal to σ^2.

2.2 CDF: Cumulative Distribution Function

The CDF of a random variable X is given by $F(x) = \Pr(X \leq x) = \sum_{x_i \leq x} f(x_i)$ in the discrete case and $F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(x)dx$ in the continuous case.

The CDF is bounded above and below by 1 and 0 respectively: $\lim_{x\to\infty} F(x) = 1$ and $\lim_{x\to-\infty} F(x) = 0$.

2.3 Expectation (or Expected Value)

The expectation or expected value of a random variable is the probability weighted average value of the random variable. Therefore, in the discrete case, we have that $E(X) = \sum_{x_i \in X} p_i x_i$ and in the continuous case we have that $E(X) = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x)dx$.

Since we sometimes deal with expected utilities, it is important not to conflate the concept of expected value and that of expected utility. It is *not* the case that the expected utility of a lottery is equal to the utility of a lottery's expected value:

$$U(L) \neq U(E(L))$$

, where $L = [p_1, ..., p_n]$ and $E(L) = \sum_{i=1}^{n} p_i x_i$.

3 Expected Utility Theorem

3.1 Assumptions on Preferences

Recall that when we started consumer theory, we placed assumptions (or restrictions) on the preferences over a set X, which denoted all of the bundle of goods (2 goods for Ec 2020a and L goods for MWG). Similarly, we are going to place certain restrictions on our preferences over the set of lotteries L. These assumptions are:

1. Rational
2. Consequentialist: for any risky alternative, only the reduced lottery over final outcomes is of relevance to the decision maker.[7] This means that our set $\mathfrak{L}$ is limited to a set of simple lotteries.
3. Continuous: $\succsim$ on the set of simple lotteries $\mathfrak{L}$ is continous if $\forall L, L', L'' \in \mathfrak{L}$, the following sets are "closed":
   (a) $\{\alpha \in [0,1] : \alpha L + (1-\alpha)L' \succsim L''\} \subset [0,1]$
   (b) $\{\alpha \in [0,1] : \alpha L + (1-\alpha)L' \precsim L''\} \subset [0,1]$.

   In words, this means that small changes in probabilities do not change the ordering between lotteries.
4. Independence Axiom ("IA"): $\succsim$ on the set of simple lotteries $\mathfrak{L}$ satisfies IA if $\forall L, L', L'' \in \mathfrak{L}$, and $\alpha \in [0,1]$, we have that:
   $L \succsim L'$ iff $\alpha L + (1-\alpha)L'' \succsim \alpha L' + (1-\alpha)L''$

3.2 Expected Utility Function

Once we assume the above, the there exists a utility function $U : \mathfrak{L} \to \mathfrak{R}$ of the following form:

$$U(L) = \sum_{i=1}^{N} p_i u_i,$$

s.t. $L \succsim L'$ iff $U(L) \geq U(L')$, where p_i is the probability of event i occurring and u_i is the utility that you get from that event. The point that needs to get across is that a lottery L is a vector of probabilities (and so the elements of L should sum to 1) that places probabilities on each of the events that is represented by the row of the vector. Then $u_i = u(x_i)$ means that the Bernoulli utility function assigns the real number u_i to the event x_i and the probability of this event happening is represented by the ith element of the lottery L.

Sometimes we write as follows (for the discrete and continuous case):

[7] MWG p. 170

$$
\begin{aligned}
E(u(x)) &= \sum_{i=1}^{N} p_i u(x_i) \\
E(u(x)) &= \int_{\underline{x}}^{\bar{x}} u(x) dF(x) = \int_{\underline{x}}^{\bar{x}} u(x) f(x) dx.
\end{aligned}
$$

Solving the UMP under uncertainty requires you to:

1. Identify each of the possible outcomes: $x_1, ..., x_K$.
2. Identify utility associated with each of the possible outcomes: $u(x_1), ..., u(x_K)$.
3. Identify probability of each outcome for each lottery: $L = \left[p_1^L, ..., p_K^L\right]$ or $F(x)$
4. Maximize by choosing the optimal outcome or the optimal lottery:
   $x^* = argmax_x E(u(x))$
   $L^* = argmax_L E(u(x))$

3.3 Transformation and Normalization of Utility

Recall the ordinality of utilities and how we transformed utility functions via a monotonic transformation. We cannot apply any type of monotone transformation when we're dealing with expected utility, as we could in consumer theory. In fact, we can only apply *positive linear* transformations to a given expected utility function. See Proposition 6.B.2 of MWG for a proof. In short, if $\tilde{U}(\cdot)$ is a monotonic transformation of $U(\cdot)$, then it must be the case that $\tilde{U}(L) = \beta U(L) + \alpha$, $\beta > 0$.

Aside from transforming the expected utility function, another simplification that might come handy in your research is the normalization of the Bernoulli utilities. Let's assume that there exists a continuum of states $x \in [\underline{x}, \bar{x}]$, then we can normalize utility such that $u(\underline{x}) = 0$ and $u(\bar{x}) = 1$.

Example 4. Given $u(x) = \sqrt{x}$; $L = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$; and $\vec{x} = [1, 4, 36]$[8], we know that $U(L) = \left(\frac{1}{3} \times \sqrt{1}\right) + \left(\frac{1}{3} \times \sqrt{4}\right) + \left(\frac{1}{3} \times \sqrt{36}\right) = 3$. If we want to transform the Bernoulli utilities such that $u(\underline{x} = 1) = 0$ and $u(\bar{x} = 36) = 1$, then:

$$
\begin{aligned}
v(x) &= \frac{u(x) - u(\underline{x})}{u(\bar{x}) - u(\underline{x})} = \frac{u(x) - 1}{6 - 1} = \frac{1}{5}\left(u(x) - 1\right) \\
v(4) &= \frac{1}{5}\left(\sqrt{4} - 1\right) = \frac{1}{5}.
\end{aligned}
$$

Therefore, we know that the new utility function$\tilde{U}(L) = p_1 v(1) + p_2 v(4) + p_3 v(36) = \frac{1}{3} \cdot \frac{1}{5} + \frac{1}{3} = \frac{2}{5}$.

[8]We can assume that the outcomes are money values.

4 Independence Axiom (or the Independence of Irrelevant Alternatives)

Definition 5. The preference relationship $\succsim$ on the set of simple lotteries $\mathfrak{L}$ satisfies IA if $\forall L, L', L'' \in \mathfrak{L}$, and $\alpha \in [0,1]$), we have that:
$L \succsim L'$ iff $\alpha L + (1-\alpha)L'' \succsim \alpha L' + (1-\alpha)L''$. In words, mixing in the same probability of a third outcome or lottery should not affect the original choice ordering (preference ordering), or if two (2) lotteries have common components, you can disregard the common component and determine the preference relationship between the two lotteries based on the non-common components.

Example 6. There are two lotteries: $L_1 = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$ and $L_2 = \left[\frac{1}{9}, \frac{7}{9}, \frac{1}{9}\right]$. Which of the two lotteries/strategies[9] do you prefer?

We use IA to simplify the two lotteries by removing the "common" parts (the L'' in the Definition above). Note that we can write the common part of the two lotteries as follows:

$$L_{common} = \left[\frac{1}{9}, \frac{1}{3}, \frac{1}{9}\right] \cdot \frac{9}{5} = \left[\frac{1}{5}, \frac{3}{5}, \frac{1}{5}\right]$$

. Now we need to find the noncommon parts of each lottery:
[10]

$$\begin{aligned} L_1^* &= \left[\frac{2}{9}, 0, \frac{2}{9}\right] \cdot \frac{9}{4} = \left[\frac{1}{2}, 0, \frac{1}{2}\right], \\ L_2^* &= \left[0, \frac{4}{9}, 0\right] \cdot \frac{9}{4} = [0, 1, 0]. \end{aligned}$$

Therefore, by IA, whether an agent prefers L_1 to L_2 will depend on whether $L_1^* \succsim L_2^*$, or whether $\frac{1}{2}u(x_1) + \frac{1}{2}u(x_2) \geq u(x_3)$.

[9]For those of you who have taken game theory, can you see why I used the term strategies?
[10]We call a lottery like L_2^* a degenerate lottery (much like a degenerate density).