Utility Maximization 3: Random Utility

Mark Dean

GR6211 - Microeconomic Analysis 1

Random Utility

- Until now, our model has been one of a decision maker who
  - Has a single, fixed utility function
  - Makes choices in order to maximize this utility function
- So if we observe the DM sometimes choose x and sometimes choose y we would declare them irrational
- But maybe this is harsh?
  - Preferences affected by some unobserved state
  - Aggregating across individuals
  - Imperfect perception leading to mistakes
- These concerns are often important when taking the model to 'real world' data

- Maybe a better model is one that accounts for this
- Random utility: Allow for random fluctuations in the utility function
- These could be due to
    - Changes in some underlying state
    - Observations from different people
    - Changes in the perception of the world

- In order to sensibly talk about this model we need to extend the data set

Definition

For a finite set X and collection of choice sets $\mathcal{D} \subset 2^X/\emptyset$ a random choice rule is a mapping $p : \mathcal{D} \to \triangle(X)$ such that $Supp(p(A)) \subset A$

- We will use $p(x, A)$ to represent the probability of choosing x from A
- Records the probability of choosing each option in each choice set
- Where does stochastic choice come from?
  - Observation from different individuals
  - Changes in choices by the same individual

Definition

A Random Utility Model (RUM) consists of a finite set of one-to-one utility functions $\mathcal{U}$ on X and a probability distribution π on $\mathcal{U}$

- Ruling out indifference (because its a pain)
- Finiteness of $\mathcal{U}$ is without loss of generality (why?)

Definition

A RUM represents a random choice rule p if, for every $A \in \mathcal{D}$

$$p(x, A) = \sum_{u \in \mathcal{U} | x = \arg\max u(A)} \pi(u)$$

- Probability of choosing x from A is equal to the probability of drawing a utility function such that x is the best thing in A
- Key feature: π does not depend on A
  - Otherwise could explain anything

- Is any choice rule compatible with RUM?
- No! One necessary condition is monotonicity

Definition

A random choice rule satisfies monotonicity if for any $x \in B \subset A \subseteq X$

$$p(x, B) \geq p(x, A)$$

- Adding alternatives to a choice set cannot increase the probability of choosing an existing option

Fact

If a Random Choice Rule is rationalizable it must satisfy monotonicity

Proof.

Follows directly from the fact that

$$\begin{aligned} & \{u \in \mathcal{U} | x = \arg\max u(A)\} \\ \subseteq \quad & \{u \in \mathcal{U} | x = \arg\max u(B)\} \end{aligned}$$

□

Rationalizing a Random Choice Rule

- So is monotonicity also sufficient for a random choice rule to be consistent with RUM?
- Unfortunately not
- Consider the following example of a stochastic choice rule on $\{x, y, z\}$

$$\begin{aligned} p(x, \{x, y\}) &= \frac{3}{4} \\ p(y, \{y, z\}) &= \frac{3}{4} \\ p(z, \{x, z\}) &= \frac{3}{4} \end{aligned}$$

- Claim: this pattern of choice is not RUM rationalizable

Rationalizing a Random Choice Rule

- Why? Well consider preference ordering such that $z \succ x$
- We know the probability of utility functions consistent with these preferences is equal to $\frac{3}{4}$
- If $z \succ x$ there are three possible linear orders

$$
\begin{aligned}
z &\succ x \succ y \\
z &\succ y \succ x \\
y &\succ z \succ x
\end{aligned}
$$

- In each case, either $y \succ x$ or $z \succ y$ or both, meaning that

$$p\left(z, \{x, z\}\right) \leq p\left(y, \{x, y\}\right) + p\left(z, \{y, z\}\right)$$

- Which is not true in this data

Characterizing Random Utility

- Do we have necessary and sufficient conditions for RUM rationalizability?
- Yes, but they are pretty horrible
- I will give you three different axioms that work
- Omit proofs, but you will play around with them a little for homework

Definition

A random choice rule satisfies the Block Marschak inequalities if for all $A \in \mathcal{D}$ and $x \in A$

$$\sum_{B|A \subseteq B} (-1)^{|B/A|} p(x, B) \geq 0$$

Theorem

A random choice rule is RUM rationalizable if and only it satisfies the Block Marschak inequalities

- Based on inclusion/exclusion restrictions for probabilities of unions of event
- Otherwise not much intuition
- Can be tested if we observe p perfectly
- Requires complete data

Axiom of Revealed Stochastic Preference

Definition

A random choice rule satisfies the Axiom of Revealed Stochastic Preference if, for any finite sequence $\{(A_1, B_1), ..., (A_n, B_n)\}$ with $A_i \in 2/\emptyset$ and $B_i \subset A_i$ (allowing for repetitions)

$$\sum_{i=1}^{n} p(B_i, A_i) \leq \max_{\succ \in \mathcal{P}} \sum_{i=1}^{n} \mathbf{1}(\succ, B_i, A_i)$$

where $\mathcal{P}$ is the set of all linear orders on X and

$$\begin{aligned} \mathbf{1}(\quad \succ \quad &, B_i, A_i) = 1 \text{ if } Max(A_i | \succ) \in B_i \\ = \quad & 0 \text{ otherwise} \end{aligned}$$

Axiom of Revealed Stochastic Preference

Theorem

A random choice rule is RUM rationalizable if and only it satisfies the Axiom of Revealed Stochastic Preference

- Does not require complete data
- Can be falsified if we observe p perfectly

Axiom of Revealed Stochastic Preference

- One way to get intuition for this is to think what it implies for deterministic choice
- Imagine that we used p to represent a deterministic choice function C, so

$$p(x, A) = 1 \text{ if } C(A) = x$$

Definition

(SARP): A choice function satisfies SARP if S (the strictly preferred relation) is acyclic

- Equivalent of GARP if there is no indifference

Axiom of Revealed Stochastic Preference

- Now imagine we had a violation of SARP so

$$x_1 S x_2 .... S x_n S x_1$$

- Implies there exists a sequence of sets $A_1 ..... A_n$ such that

$$\begin{aligned} x_i &\in C(A_i) \text{ and } x_{i+1} \in A_i \text{ for } i < n \\ x_n &\in C(A_n) \text{ and } x_1 \in A_n \end{aligned}$$

- So consider the sequence $\{(x_i, A_i)\}_{i=1}^n$
- We know that

$$\sum_{i=1}^{n} p(x_i, A_i) = n$$

- But we also know that this data can't be rationalized by any preference relation, so

$$\max_{\succ \in \mathcal{P}} \sum_{i=1}^{n} \mathbf{1}(\succ, x_i, A_i) < n$$

- So ASRP implies SARP

- Consider a data set consisting of choices from $\{a_1, a_2\}$, $\{a_1, a_2, a_3\}$ and $\{a_1, a_2, a_3, a_4\}$
- Construct vectors each entry of which relates to a given choice from each choice set

$$\begin{matrix} a_1 | \{a_1, a_2\} \\ a_2 | \{a_1, a_2\} \\ a_1 | \{a_1, a_2, a_3\} \\ a_2 | \{a_1, a_2, a_3\} \\ a_3 | \{a_1, a_2, a_3\} \\ a_1 | \{a_1, a_2, a_3, a_4\} \\ a_2 | \{a_1, a_2, a_3, a_4\} \\ a_3 | \{a_1, a_2, a_3, a_4\} \\ a_4 | \{a_1, a_2, a_3, a_4\} \end{matrix}$$

- Construct a matrix of all possible rationalizable choice vectors

$$
\begin{array}{c}
a_1 | \{a_1, a_2\} \\
a_2 | \{a_1, a_2\} \\
a_1 | \{a_1, a_2, a_3\} \\
a_2 | \{a_1, a_2, a_3\} \\
a_3 | \{a_1, a_2, a_3\} \\
a_1 | \{a_1, a_2, a_3, a_4\} \\
a_2 | \{a_1, a_2, a_3, a_4\} \\
a_3 | \{a_1, a_2, a_3, a_4\} \\
a_4 | \{a_1, a_2, a_3, a_4\}
\end{array}
\quad
\left\{
\begin{array}{cccc}
1 & 1 & 0 & \\
0 & 0 & 1 & \\
1 & 1 & 0 & \\
0 & 0 & 0 & \\
0 & 0 & 1 & .... \\
1 & 0 & 0 & \\
0 & 0 & 0 & \\
0 & 0 & 1 & \\
0 & 1 & 0 &
\end{array}
\right\} = A
$$

- Let P be the observed choice probabilities associated with each row of the matrix A

Theorem

P is rationalizable by RUM if and only if their exists a probability vector v such that

$$Av = P$$

- Obviously true, but doesn't offer much insight
- Computationally feasible
- Kitamura Stoye offer a statistical test even if we only observe estimates of p

The Luce Model

- Random utility is a very interesting model in principle
- But its full generality it may not be very useful
  - Predictions are weak
  - Axiomatization doesn't provide much intuition
- In practice it may be more useful to work with specific models in the random utility class

- One particularly popular version is the Luce model

Definition

A Random Choice rule on a finite set X has a Luce representation if there exists a utility function $u : X \rightarrow \mathbb{R}_{++}$ such that for every $A \in \mathcal{D}$ and $x \in A$

$$p(x, A) = \frac{u(x)}{\sum_{y \in A} u(y)}$$

- Advantages:
  - Captures the intuitive notion that 'better things are chosen more often'
  - Equivalent to the Logit form where choice is based on v given by

    $$v(x) = u(x) + \varepsilon$$

    and ε has an extreme value type 1 distribution
  - Extremely heavily used in applied work

Extension 2: Luce

- The Luce model also has a very clean axiomatization

Definition

A random choice rule p on a set X satisfies stochastic independence of irrelevant alternatives if and only if, for any $x, y \in X$ and $A, B \in \mathcal{D}$ such that $x, y \in A \cap B$

$$\frac{p(x, A)}{p(y, A)} = \frac{p(x, B)}{p(y, B)}$$

Theorem

A random choice rule is rationalizable by the Luce model if and only if it satisfies Stochastic IIA

- Problem: Stochastic IIA sometimes not very appealing:
  - Consider {red bus, car} vs {red bus, blue bus, car}