9. Stochastic Choice

Duarte Gonçalves

University College London

MRes Microconomics

Overview

Stochastic choice: Population-level data, choice frequencies; Individual-level data.

Why is individual choice (seemingly) random?

- **Mistakes:** trembling hands, mistakes.

February 2022, Spain

Spain's Socialist-led government was trying to get a significant labour reform approved in parliament.

While the labour reform was agreed with business and union organisations in December following months of negotiations, the government was unable to build a stable parliamentary majority for Thursday's vote.

The government had previously secured the support of the two MPs from UPN and expected them to vote "yes" in the vote; they declared throughout the afternoon of the previous day and the day of the vote that they would vote in favor of the reform.

However, backstage negotiations with the leader of the opposition would see these two MPs vote "no" leading to a rejection of the reform with 174 votes in favor and 175 against!

However... an MP from the opposition, Alberto Casero Ávila, makes a mistake when casting his vote and voted "yes". The final count allowed the necessary majority to be reached and the reform passed.

Overview

Stochastic choice: Population-level data, choice frequencies; Individual-level data.

Why is individual choice (seemingly) random?

- **Mistakes:** trembling hands, mistakes.
- **Information:** Changes in the environment; agent doesn't see choice as random. Information arrival may affect how individual's compare alternatives: at the restaurant you are told that sea breams are particularly good that day.
- **Randomly fluctuating tastes:** these can be due to alternating diversity (I like both meat and fish, but I alternate between the two instead of buying half a portion of each every time) or some other factor that is unobserved by the analyst.
- **Random attention, misperception:** items considered change (e.g., you fail to find your preferred jam in a supermarket so you go for another one; the analyst sees the jam in stock and changes in behaviour are seen as stochastic).
- **Experimentation:** you moved countries and you are trying to find out which kind of bread you like the most at your new local bakery.
- Information, random attention, information acquisition, experimentation: topics developed in Term 2 topics course
- **Today:** A bird's-eye view of stochastic choice.

Overview

1. Stochastic Choice
2. Stochastic Choice and Random Utility
3. Connecting Stochastic and Deterministic Choice
4. Discrete Choice
5. Controlled Randomisation
6. More

Overview

Stochastic Choice

Primitives

Alternatives: X (assume finite for convenience).

Menus: $\mathcal{A} := \{A \subseteq X | A \neq \emptyset\}$.

Shift focus to **stochastic choice**: from choice functions to choice frequencies.

Recall choice function: $C(A) \subseteq A, A \neq \emptyset \implies C(A) \neq \emptyset$.

Definition

$\rho : \mathcal{A} \to \Delta(X)$ is a **stochastic choice function (SCF)** iff $\text{supp}(\rho(A)) \subseteq A$.

$\rho(x, A) = \rho(x|A)$: prob. choosing x from A.

ρ (instead of C) describes observable data.

WT obtain useful characterisation of ρ

Random Utility

Definition

A **random utility model (RUM)** is a pair $(\mathcal{U}, \pi)$ s.t. $\mathcal{U} := \{u : X \to \mathbb{R}\} \subseteq \mathbb{R}^X$, and $\pi \in \Delta(\mathcal{U})$.

Stochastic choice because population heterogeneous or varying preferences.

Menu A drawn at random

Preference drawn at random

- Preference as individual (or mood, shock, etc.)
- Independent of menu; hard with heterogenous population assumption: need that different people choose from the same menu with the same prob.

Random Utility

Definition

A stochastic choice function ρ admits a **random utility (RU) representation** iff there is a random utility model $(\mathcal{U}, \pi)$ s.t. $\rho(x, A) = \pi(\{u \in \mathcal{U} \mid x \in \arg\max_{y \in A} u(y)\})$.

Key assumption: π invariant wrt feasible set A, ow RUM has no empirical content (Why?)

WLOG, can replace $\mathcal{U}$ with the set of all preference relations on X,

$\tilde{R} := \{\succsim \in X^2 | \succsim$ s.t. completeness and transitivity$\}$.

(Possibly better, as when X is finite $\tilde{R}$ is finite but $\mathbb{R}^X$ is not.)

NB: WLOG to focus on $\succsim$ s.t. $\nexists x \neq y : x \sim y$

$R := \{\succ \in \tilde{R} | \forall x, y \in X : x \neq y, \neg(x \sim y)\}$.

Proposition

ρ has a RU representation if and only if there is $\pi \in \Delta(R) : \rho(x, A) = \pi(\{\succsim \in R | x = \arg\max_{\succsim} A\})$.

We'll call both $\pi \in \Delta(R)$ and $\pi \in \Delta(\mathcal{U})$ RUM

Random Utility

Definition

A stochastic choice function ρ sat. **monotonicity** if $\forall x \in B \subseteq A \subseteq X$, $\rho(x, B) \geq \rho(x, A)$.

Also called 'regularity' (too presumptuous and uninformative an expression).

Proposition

ρ has a RU representation only if it sat. monotonicity.

Proof

$\forall x \in B \subseteq A, \forall \succsim \in R, x = \arg\max_{\succsim} A \implies x = \arg\max_{\succsim} B.$

$\implies \forall x \in B \subseteq A, \{\succsim \in R | x \in \arg\max_{\succsim} A\} \subseteq \{\succsim \in R | x \in \arg\max_{\succsim} B\}.$

$\implies \forall x \in B \subseteq A, \rho(x, A) = \pi(\{\succsim \in R | x \in \arg\max_{\succsim} A\}) \leq \pi(\{\succsim \in R | x \in \arg\max_{\succsim} B\}) = \rho(x, B).$

Random Utility

Definition

A stochastic choice function ρ sat. **monotonicity** if $\forall x \in B \subseteq A \subseteq X$, $\rho(x, B) \geq \rho(x, A)$.

Plenty of reasons why SCF monotonicity may fail.

- With larger choice sets, it may be more difficult to find and compare items (search is costly, and so is thinking about the differences and assessing alternatives!). DM may end up choosing a particularly salient but worse product more often than with a smaller choice set. Status quo.
- Decoy effect: Economist.com subscription $59, Print subscription $125, Print + Economist.com subscription $125.

Random Utility

Definition

A stochastic choice function ρ sat. **monotonicity** if $\forall x \in B \subseteq A \subseteq X$, $\rho(x,B) \geq \rho(x,A)$.

Is monotonicity sufficient for SCF to have RU representation?

Proposition

If $|X| \leq 3$, then ρ sat. monotonicity if and only if it admits RU representation.

Proof

$X = \{a, b, c\}$. $|R| = 6$. Identify $\succsim \in R$ with preference ordering $(a, b, c) \implies a \succ b \succ c$.

Note: (a) $\pi((a,b,c)) + \pi((a,c,b)) = \rho(a,X)$, (b) $\pi((b,a,c)) + \pi((b,c,a)) = \rho(b,X)$,
(c) $\pi((c,b,a)) + \pi((c,a,b)) = \rho(c,X)$

Then: $\rho(a,\{a,b\}) = \rho(a,X) + \pi((c,a,b)) \iff \pi((c,a,b)) = \rho(a,\{a,b\}) - \rho(a,X)$

By similar logic:

- $\pi((c,a,b)) = \rho(a,\{a,b\}) - \rho(a,X)$
- $\pi((c,b,a)) = \rho(b,\{a,b\}) - \rho(b,X)$
- $\pi((b,a,c)) = \rho(a,\{a,c\}) - \rho(a,X)$
- $\pi((b,c,a)) = \rho(c,\{a,c\}) - \rho(c,X)$
- $\pi((a,b,c)) = \rho(b,\{b,c\}) - \rho(b,X)$
- $\pi((a,c,b)) = \rho(b,\{b,c\}) - \rho(c,X)$

$\therefore \pi \geq 0 \iff \rho$ sat. monotonicity. By definition, $\pi(R) = 1$. $\square$

Random Utility

Beyond 3 alternatives, monotonicity is not sufficient.

Directly generalise difference idea from proof with 3 items.

Definition (Block & Marschak 1960)

SCF ρ sat. Block-Marschak inequalities iff $\forall x \in X, A \in \mathcal{A}$, $\sum_{A \subseteq B} (-1)^{|B \setminus A|} \rho(x, B) \geq 0$.

Definition (McFadden & Richter 1990)

A SCF ρ sat. the **Axiom of Revealed Stochastic Preference** (ARSP) iff, $\forall$ finite sequences $\{(A_1, B_1), (A_2, B_2), \cdots, (A_n, B_n)\}$ with $A_i \in \mathcal{A}$ and $B_i \subseteq A_i$ (allowing for repetitions) $\sum_{i=1}^{n} \rho(B_i, A_i) \leq \max_{\succsim \in R} \sum_{i=1}^{n} 1\{\arg\max_{\succsim} A_i \in B_i\}$.

Random Utility

Theorem

The following are equivalent:

(i) SCF ρ admits a RU representation.

(ii) SCF ρ sat. Block-Marschak inequalities.

(iii) SCF ρ sat. ARSP.

Not quite very intuitive? A connection with deterministic choice and SARP.

Still, enables tests of RUM: Kitamura & Stoye (2018 Ecta)

Overview

1. Stochastic Choice
2. Stochastic Choice and Random Utility
3. Connecting Stochastic and Deterministic Choice
4. Discrete Choice
5. Controlled Randomisation
6. More

Connecting with SARP

Recall: choice function $C : \mathcal{A} \to \mathcal{A}$, x is directly revealed strictly preferred to y iff $\exists A \in \mathcal{A}$ s.t. $x \in C(A)$ and $y \in A \setminus C(A)$.

Fixing C, let $S \subseteq X^2$ s.t. $xSy \iff x$ is directly revealed strictly preferred to y.

Definition

A choice function $C : \mathcal{A} \to \mathcal{A}$ sat. the **strong axiom of revealed preference** (SARP) iff there is no sequence $\{x_0, x_1, ..., x_n\} \subseteq X$ s.t. x_i is directly revealed strictly preferred to $x_{i+1 \bmod (n+1)}$ for $i = 0, ..., n$.

Proposition

Let X be finite and C be a singleton-valued choice function on X, i.e., $C : \mathcal{A} \to \mathcal{A}$, $\mathcal{A} = 2^X \setminus \{\emptyset\}$.

(i) $\exists$ a preference relation $\succsim$ on X s.t. $C(A) = \arg\max_{\succsim} A \forall A \in \mathcal{A}$ if and only if C satisfies SARP.

(ii) Furthermore, any such $\succsim$ is s.t. $\forall x, y \in X : x \neq y, x \succsim y \implies \neg(y \succsim x)$.

Connecting with SARP

Proposition

Let X be finite and C be a singleton-valued choice function on X, i.e., $C : \mathcal{A} \to \mathcal{A}$, $\mathcal{A} = 2^X \setminus \{\emptyset\}$.

(i) $\exists$ a preference relation $\succsim$ on X s.t. $C(A) = \arg\max_{\succsim} A \forall A \in \mathcal{A}$ if and only if C satisfies SARP.

(ii) Furthermore, any such $\succsim$ is s.t. $\forall x, y \in X : x \neq y, x \succsim y \implies \neg(y \succsim x)$.

Proof

If xSy, then for some $A \in \mathcal{A}$, $x \in \arg\max_{\succsim} A$ and $y \notin \arg\max_{\succsim} A$
$\implies x \succsim y$ and $\neg(y \succsim x) \implies x \succ y$.

(i) $\implies$: Suppose that SARP is violated but $\exists$ preference relation $\succsim \in X^2$ that rationalised C.

Then, $\exists \{x_i\}_{i=0,\ldots,n} : x_0 \succ x_1 \succ \cdots \succ x_n \succ x_0$, which contradicts the fact that if $\succsim$ is transitive then so is its strict part $\succ$.

Connecting with SARP

Proposition

Let X be finite and C be a singleton-valued choice function on X, i.e., $C : \mathcal{A} \to \mathcal{A}$, $\mathcal{A} = 2^X \setminus \{\emptyset\}$.

(i) $\exists$ a preference relation $\succsim$ on X s.t. $C(A) = \arg\max_{\succsim} A \forall A \in \mathcal{A}$ if and only if C satisfies SARP.

(ii) Furthermore, any such $\succsim$ is s.t. $\forall x, y \in X : x \neq y, x \succsim y \implies \neg(y \succsim x)$.

Proof

(i) $\Longleftarrow$: Suppose that SARP is satisfied. Define $\succsim$ s.t. $x \succsim y \iff \exists A \in \mathcal{A} : x, y \in A$ and $x = C(A)$.

By assumption C is singleton-value, and $\forall x, y \in X$, $x = C(\{x, y\})$ or $y = C(\{x, y\})$, which implies completeness of $\succsim$.

SARP immediately implies that $\succsim$ will be transitive.

(ii) Immediate from the singleton-valuedness and nonemptiness of C together with there being a preference relation $\succsim$ on X s.t. $C(A) = \arg\max_{\succsim} A \forall A \in \mathcal{A}$. □

Connecting with SARP

What goes wrong if C is not singleton-valued?

Let $X = \{x, y, z\}$, $C(\{x, y, z\}) = \{x, z\}$; $C(\{x, y\}) = \{x\}$; $C(\{y, z\}) = \{y, z\}$.

C does not violate SARP, but it violates HARP, hence there is no preference relation $\succsim$ on X that can rationalise C.

Axiom of Revealed Stochastic Preference

Proposition

Let ρ be a degenerate SCF and $C : \rho(x, A) = 1 \implies C(A) = x$. C sat. SARP $\iff$ ρ sat. ARSP.

Proof

(1) Only if: C violates SARP $\implies$ ρ violates ARSP.

If SARP is violated, $\exists$ sequence $\{A_i\}_{i=0,\dots,n}$, such that $C(A_{i+1 \bmod (n+1)}) \in A_i \setminus C(A_i)$.

Then, $\sum_{i=0}^{n} \mathbf{1}\{C(A_i) \in \{C(A_i)\}\} = n + 1$.

Further, we can take the violating sequence so that each choice $C(A_i)$ is distinct. (Prove it!)

Violation of SARP $\implies$ data cannot be rationalised by any preference relation and $\max_{\succsim \in R} \sum_{i=1}^{n} \mathbf{1}\{\arg\max_{\succsim} A_i \in B_i\} < n + 1$.

Axiom of Revealed Stochastic Preference

Proposition

Let ρ be a degenerate SCF and $C : \rho(x, A) = 1 \implies C(A) = x$. C sat. SARP $\iff$ ρ sat. ARSP.

Proof

(1) Only if: C violates SARP $\implies$ ρ violates ARSP.

(2) If: C sat. SARP $\implies$ ρ sat. ARSP.

Let $\succsim$ rationalise C. Then, $\succsim \in R$.

$$\sum_{i=1}^{n} \mathbf{1}\{C(A_i) \in B_i\} = \sum_{i=1}^{n} \mathbf{1}\{\arg\max_{\succsim} A_i \in B_i\} \leq \max_{\succsim \in R} \sum_{i=1}^{n} \mathbf{1}\{\arg\max_{\succsim} A_i \in B_i\}.$$

□

Overview

1. Stochastic Choice

2. Stochastic Choice and Random Utility

3. Connecting Stochastic and Deterministic Choice

4. Discrete Choice
   - Luce Model
   - Blue Bus/Red Bus
   - SCP for Random Utility

5. Controlled Randomisation

6. More

Discrete Choice

Definition

A stochastic choice function ρ admits a **discrete choice (DC) representation** iff $\exists v : X \to \mathbb{R}$ and, $\forall x \in X$, a random variable $\varepsilon(x)$ with full support on the real line s.t. $\rho(x, A) = \mathbb{P}(x = \arg\max_{y \in A} v(y) + \varepsilon(y))$.

Interpretation: $u(x) = v(x) + \varepsilon(x)$.

- v is deterministic utility function
- $\varepsilon(x)$ is a 'utility shock' to $v(x)$.
- Sometimes called additive RUM.

Discrete Choice

Definition

A stochastic choice function ρ sat. **positivity** iff $\rho(x, A) > 0$ $\forall x \in A$ and $\forall A \in \mathcal{A}$.

Necessary for DC. Also: cannot be falsified with finite data.

Theorem

Let X be finite and ρ be a SCF sat. positivity. ρ admits a RU representation if and only if ρ admits a DC representation.

Discrete Choice

Note that without further assumption, $\rho(x, \{x, y\}) > 1/2$ **does not imply nor is implied by** $v(x) > v(y)$...

Definition

A stochastic choice function ρ admits an **iid discrete choice representation** iff $\exists v : X \to \mathbb{R}$ and, $\forall x \in X$, a random variable $\varepsilon(x)$ with full support on the real line s.t. $\rho(x, A) = \mathbb{P}(x = \arg\max_{y \in A} v(y) + \varepsilon(y))$ and $\varepsilon(x)$ are iid $\forall x \in X$.

Cannot identify v without further assumptions (manipulating v, observing time – more later).

Assume $\varepsilon(x)$ iid: **iid discrete choice representation**.

Discrete Choice

Observation

(v, ε) is an iid DC representation of SCF ρ if and only if, $\forall \alpha > 0, \beta, \gamma \in \mathbb{R}$, $(\alpha v + \beta, \alpha\varepsilon + \gamma)$ is an iid DC representation of SCF ρ.

Special familiar cases:

- logit: $\varepsilon(x)$ follows a zero mean extreme value distribution
- probit: $\varepsilon(x)$ follows a zero mean Normal distribution

iid DC: let $\varepsilon_{x,y} := \varepsilon(x) - \varepsilon(y) \sim F(\cdot)$.

Then: $\rho(x, \{x, y\}) = \mathbb{P}(\varepsilon(y) - \varepsilon(x) \leq v(x) - v(y)) = F(v(x) - v(y))$.

Binary Discrete Choice

Definition

A stochastic choice function ρ admits a **Fechnerian representation** iff $\exists v : X \to \mathbb{R}$ and strictly increasing $F : \mathbb{R} \to [0,1]$ s.t. $\rho(x, \{x, y\}) = F(v(x) - v(y))\ \forall x, y \in X$.

RUM and Fechnerian models non-nested.

Fechnerian models sat. **Weak Stochastic Transitivity**:

$\succsim \subseteq X^2 : x \succsim y \iff \rho(x, \{x, y\}) \geq 1/2$ is transitive.

Fechnerian models are special case of simple scalability models:

Definition

A stochastic choice function ρ admits a **simple scalability representation** iff $\exists v : X \to \mathbb{R}$ and strictly increasing $F : \mathbb{R}^2 \to [0,1]$ s.t. $\rho(x, \{x, y\}) = F(v(x), v(y))\ \forall x, y \in X$.

Binary Discrete Choice

Definition

Let ρ be a SCF on X and let x, y, z be s.t. $\rho(x,\{x,y\}), \rho(y,\{y,z\}) \geq 1/2$.$\rho$ sat.

(i) **Weak Stochastic Transitivity** iff $\rho(x,\{x,z\}) \geq 1/2$;

(ii) **Strong Stochastic Transitivity** iff $\rho(x,\{x,z\}) \geq \max\{\rho(x,\{x,y\}), \rho(y,\{y,z\})\}$;

(iii) **Tversky-Russo Independence** iff $\forall x, y, w, z \in X$, $\rho(x,\{x,w\}) \geq \rho(y,\{y,w\}) \iff \rho(x,\{x,z\}) \geq \rho(y,\{y,z\})$;

(iv) **Tversky-Russo Substitutability** iff $\forall x, y, z \in X$, $\rho(x,\{x,z\}) \geq \rho(y,\{y,z\}) \iff \rho(x,\{x,y\}) \geq 1/2$.

Theorem (Tversky & Russo (1969 JMathPsy))

The following are equivalent:

(i) ρ sat. strong stochastic transitivity;

(ii) ρ sat. TR independence;

(iii) ρ sat. TR substitutability;

(iv) ρ admits a simple scalability representation.

Luce Model

Definition

A stochastic choice function ρ admits a **Luce representation** iff $\exists v : X \to \mathbb{R}_{++}$ s.t. $\rho(x, A) := \frac{v(x)}{\sum_{y \in A} v(y)}$.

Interpretation: $v(x)$ as intensity of preference for x; choice prob. $\propto$ preference intensity.

Arguably bread-and-butter of much empirical and structural work.

Theorem (McFadden (1973))

The following are equivalent:

(i) ρ admits a logit representation with v;
(iid DC with $\varepsilon(x)$ $\sim$zero mean extreme value distribution)

(ii) ρ admits a Luce representation with $\tilde{v} = \exp \circ v$.

Luce Model

Some properties of Luce/logit representation:

Definition

A stochastic choice function ρ sat.

Luce's independence of irrelevant alternatives iff $\forall x, y \in A \cap B$, whenever probabilities are positive, $\dfrac{\rho(x,A)}{\rho(y,A)} = \dfrac{\rho(x,B)}{\rho(y,B)}$;

Luce's choice property iff $\forall x \in B \subseteq A$, $\rho(x,A) = \rho(x,A)\rho(B,A)$, where $\rho(B,A) := \sum_{y \in B} \rho(y,A)$.

Turns out these pin-down a Luce representation!

Luce Model

Theorem (Luce (1969))

Let X be finite and ρ a SCF on X. The following are equivalent:

(i) ρ satisfies positivity and Luce's IIA;

(ii) ρ satisfies positivity and Luce's choice property;

(iii) ρ admits a Luce representation.

Characterisation (i) allows you to test if your data is consistent with logit choice (rather than just assuming it);
(ii) provides useful properties that you can use in derivations.

Luce Model

Theorem (Luce (1969))

(i) positivity and Luce's IIA $\iff$ (iii) Luce representation.

Proof

- (i) positivity and Luce's IIA $\iff$ (iii) Luce representation.
  Focus on $\implies$.

  Define $v(x) := \rho(x, X)$ and fix $x \in A$.

$$\frac{\rho(y,A)}{\rho(x,A)} = \frac{\rho(y,X)}{\rho(x,X)} \iff \rho(y,A) = \rho(y,X)\frac{\rho(x,A)}{\rho(x,X)} = \frac{v(y)}{v(x)}\rho(x,A)$$

$$\implies 1 = \sum_{y \in A} \rho(y,A) = \sum_{y \in A} \frac{v(y)}{v(x)}\rho(x,A)$$

$$\iff \rho(x,A) = \frac{v(x)}{\sum_{y \in A} v(y)}.$$

Luce Model

Theorem (Luce (1969))

(i) positivity and Luce's IIA $\iff$ (iii) Luce representation.

Proof

- (i) positivity and Luce's choice property $\iff$ (iii) Luce representation.
  Again focus on $\implies$.

  Define v in same manner.

$$\rho(x, X) = \rho(x, A)\rho(A, X) = \rho(x, A) \sum_{y \in A} \rho(y, X) \iff \rho(x, A) = \frac{v(x)}{\sum_{y \in A} v(y)}.$$

Blue Bus/Red Bus (Debreu's Critique)

Suppose DM chooses between taking red buses (rb), blue buses (bb), and trains (t).

Suppose we observe that $\rho(t, \{rb, t\}) = \rho(t, \{bb, t\}) = \rho(rb, \{rb, bb\}) = 1/2$.
$\rho(t, \{rb, t\}) = \rho(rb, \{rb, t\}) = 1/2$.

If ρ admits an iid discrete choice representation, it must be that $\rho(t, \{t, rb, bb\}) = 1/3$.

Issue: if we had n colors of buses, we would then have that $\rho(t, X) = 1/(n+1)$, which makes no sense if DM does not care for the color of the bus. How to handle this? Use RUM: place equal prob on $rb \succ bb \succ t$, $bb \succ rb \succ t$, $t \succ rb \succ bb$, and $t \succ bb \succ rb$

or use parametric discrete choice families (e.g., nested logit: color as an attribute of buses).

Controlled Randomisation

Definition

Let $V(\rho) := \sum_{x \in X} \rho(x)v(x) - C(\rho)$, where $C : \Delta(X) \to \mathbb{R}$, $v : X \to \mathbb{R}$. SCF ρ on X admits a **perturbed utility representation** if $\rho = \arg\max_{r \in \Delta(X)} V(r)$.

Moreover, ρ admits an **additive perturbed utility representation** if, for any $A \subseteq X$ and some $\alpha > 0$, $\beta \in R$, $C(\rho) = \alpha \sum_{x \in A} c(\rho(x)) + \beta$, with $c : [0, 1] \to \mathbb{R} \cup \{\infty\}$.

Interpretation:

- Trembling hands with implementation costs
- Cost to pay attention, be precise
- Hedging against ambiguity
- Regret minimisation

Examples of cost functions:

- Log: $c(x) = -\ln x$
- Quadratic: $c(x) = x^2$
- Entropy: $c(x) = x \ln x$

Controlled Randomisation

Proposition (Anderson, de Palma, & Thisse (1992))

Let X be finite and ρ a SCF on X. The following are equivalent:

(i) ρ admits a Luce representation;

(ii) ρ admits an additive perturbed utility representation with entropy costs.

Proof Sketch

Note that $\rho(x, A) > 0 \; \forall x \in A$ (ow infinite marginal cost).

Lagrangian: $\sum_{x \in A} \rho(x, A) v(x) - \rho(x, A) \ln(\rho(x, A)) + \lambda(1 - \sum_{x \in A} \rho(x, A))$.

FOC: $v(x) - \ln \rho(x, A) = \lambda + 1 \; \forall x \in A \implies v(y) - v(x) = \ln \rho(y, A)/\rho(x, A)$

$\iff \rho(y) = \rho(x) \exp(v(y) - v(x))$.

$\implies 1 = \sum_{y \in A} \rho(y, A) = \rho(x, A) \sum_{y \in A} \exp(v(y) - v(x))$

$$\iff \rho(x, A) = \frac{\exp(v(x))}{\sum_{y \in A} \exp(v(y))}.$$

Controlled Randomisation

Definition

Let X be finite and ρ a SCF on X. ρ sat.

(i) **ordinal independence of irrelevant alternatives** iff $\exists \phi : (0,1) \to \mathbb{R}_+$ s.t. $\forall x, y \in A \cap B$, whenever probabilities are positive, $\frac{\phi(\rho(x,A))}{\phi(\rho(y,A))} = \frac{\phi(\rho(x,B))}{\phi(\rho(y,B))}$;

(ii) **acyclicity** iff for any permutations $\pi, \tilde{\pi}$ on $[n]$, whenever $\rho(x_1, A_1) > \rho(x_{pi(1)}, A_{\tilde{\pi}(1)})$ and $\rho(x_k, A_k) \geq \rho(x_{pi(k)}, A_{\tilde{\pi}(k)})$ for any $1 < k < n$, it is the case that $\rho(x_n, A_n) < \rho(x_{pi(n)}, A_{\tilde{\pi}(n)})$.

Ordinal-IIA as a generalisation of Luce's IIA, where ϕ =id

Controlled Randomisation

Theorem (Fudenberg, Iijima, & Strzalecki (2015 Ecta))

Let X be finite and ρ a SCF on X. The following are equivalent:

(i) ρ admits an additive perturbed utility representation such that c is $\mathcal{C}^1$, strictly convex and $c'(0^+) = -\infty$;

(ii) ρ satisfies ordinal IIA;

(iii) ρ satisfies acyclicity.

Proof (i) $\Longleftrightarrow$ (ii) similar to Luce's (using FOC).

In paper: characterisation of menu-dependent costs, comparison with RUM (non-nested).

Overview

Learning and Information Acquisition

Example (Luce & Raiffa, 1957)

A person enters a new restaurant. The waiter informs that that evening there is the chicken and the steak tartare.

In a first-rate restaurant, the DM's preferred alternative would've been the tartare, but considering the unknown surroundings, the DM elects the chicken.

Soon after the waiter returns from the kitchen, apologizes profusely, blaming the uncommunicative chef for forgetting to say that frogs' legs are also on the menu.

The DM dislikes frogs' legs and would always prefer chicken, yet their response is "Splendid, I'll change my order to steak tartare".

Learning and Information Acquisition

Randomness in information/perception $\implies$ Randomness in choice.

Also: agency! DM often can choose what to pay attention to or what to learn about.

Examples?

Learning and paying attention is costly! Cost-benefit analysis in information.

Requires understanding:

(i) what is information,

(ii) what's the value of information in a problem,

(iii) how to model cost of information.

See Strzalecki's slides. More in 2nd Year!

More

Costly Information Acquisition: Raiffa & Schlaifer (1961), Sims (2003 JME), Matejka & McKay (2015 AER), Caplin & Dean (2015 AER)

Consideration Sets and Attention: Random attention filters: Cattaneo, Ma, Masatlioglu, & Suleymanov (2020 JPE)

Sequential Sampling and Timed Stochastic Choice: Fudenberg, Strack, & Strzalecki (2018 AER), Alós-Ferrer, Fehr, & Netzer (2022 JPE), Gonçalves (2024 WP)