Ch.3 More Models and Properties

Chang-Yeop Lee

The Ohio State University

June 4, 2024

Review

Goal: Rationalize random choice data $\rho(x, A)$

- Pref-based: Attach μ on each pref orderings
- Random utility: Add randomness on utility values

Review

Goal: Rationalize random choice data $\rho(x, A)$

- Pref-based: Attach μ on each pref orderings
- Random utility: Add randomness on utility values

Proposition 1.9, 1.12 Followings are equivalent for finite X:

- ρ is represented by a distribution over preferences
- $\rho \sim RU$
- $\rho \sim ARU$

Review

Goal: Rationalize random choice data $\rho(x, A)$

- Pref-based: Attach μ on each pref orderings
- Random utility: Add randomness on utility values

Proposition 1.9, 1.12 Followings are equivalent for finite X:

- ρ is represented by a distribution over preferences
- $\rho \sim RU$
- $\rho \sim ARU$

Main axioms of $\rho \sim$ RU (of finite X)

- Regularity: Necessary but limitedly sufficient ($|X| \leq 3$)
- Supermodularity: Additional property if $|X| = 4$
- Block-Marshack: Necessary and sufficient

Preview

Takeaway – Add more structure on RU (parameterize)

Main questions on parametric approaches

1. Characterization: Which data patterns that model allows?
2. Identification: Unique data interpretations?
3. Comparative Statics: Meaning of parameter values?
   ex.) risk preference, inattention

Preview

Takeaway – Add more structure on RU (parameterize)

Main questions on parametric approaches

1. Characterization: Which data patterns that model allows?
2. Identification: Unique data interpretations?
3. Comparative Statics: Meaning of parameter values?
   ex.) risk preference, inattention

This week: Study ARU model

Outline

1. *i.i.d.* ARU
2. Fechnerian Model
3. Logit Model
   3-1. Logit (or Luce) Model
   3-2. Mixed Logit
   3-3. Nested Logit
   3-4. Perturbed Utility

i.i.d. ARU

Definition (*i.i.d* ARU) $\rho \sim$ *i.i.d.* $ARU(v, \epsilon)$ if $\rho \sim ARU(v, \epsilon)$ where for all $x, y \in X$, $\epsilon(x)$ and $\epsilon(y)$ are *i.i.d.*.

i.i.d. ARU

Definition (*i.i.d* ARU) $\rho \sim i.i.d.\ ARU(v, \epsilon)$ if $\rho \sim ARU(v, \epsilon)$ where for all $x, y \in X$, $\epsilon(x)$ and $\epsilon(y)$ are *i.i.d.*.

Examples: Logit (T1EV), Probit (Normal)

i.i.d. ARU

> **Definition (*i.i.d* ARU)** $\rho \sim i.i.d.\ ARU(v, \epsilon)$ if $\rho \sim ARU(v, \epsilon)$ where for all $x, y \in X$, $\epsilon(x)$ and $\epsilon(y)$ are *i.i.d.*.

Examples: Logit (T1EV), Probit (Normal)

Binary choice: If $\rho \sim i.i.d.\ ARU$, then $\exists v$ and F such that

$$\begin{aligned}\rho(x, y) &= \mathbb{P}(v(x) + \epsilon(x) \geq v(y) + \epsilon(y)) \\ &= \mathbb{P}(\epsilon(y) - \epsilon(x) \leq v(x) - v(y)) = F(v(x) - v(y))\end{aligned}$$

where $F(\cdot)$: CDF of $\eta \equiv \epsilon(y) - \epsilon(x)$ (Fechnerian)

$|X| > 2$? Not trivial – $\rho(x, A) = \mathbb{E}_{\epsilon(x)}[\prod_{y \in A \setminus \{x\}} \Phi(v(x) - v(y) + \epsilon(x))]$

Properties of *i.i.d.* ARU

Proposition 3.10 Suppose ρ has two i.i.d representation with positive density $(v_1, \epsilon_1), (v_2, \epsilon_2)$ and let F_i be a cdf of $\epsilon_1 - \epsilon_2$. If the range of v_1 is a nontrivial interval, then there exists $\alpha > 0$ and β such that

$$v_2 = \alpha v_1 + \beta, \text{ and}$$
$$F_2(t) = F_1(\alpha^{-1} t), \text{ for all } t \in \{v_2(x) - v_2(y) | x, y \in X\}$$

Implication: v has a unique representation in positive **affine** transformation

- ϵ distribution may not be unique (but diff. in ϵ is!)

Violation of **i.i.d.**

Blue Bus and Red Bus Suppose $X = \{t, bb, rb\}$ and $\rho(t, bb) = \rho(t, rb) = \rho(bb, rb) = 1/2$.

Violation of **i.i.d.**

Blue Bus and Red Bus Suppose $X = \{t, bb, rb\}$ and $\rho(t, bb) = \rho(t, rb) = \rho(bb, rb) = 1/2$.

i.i.d. ARU ρ represented by $v(t) = v(bb) = v(rb)$
Predicts $\rho(t, X) = \rho(bb, X) = \rho(rb, X) = 1/3$

Violation of **i.i.d.**

Blue Bus and Red Bus Suppose $X = \{t, bb, rb\}$ and $\rho(t, bb) = \rho(t, rb) = \rho(bb, rb) = 1/2$.

i.i.d. ARU ρ represented by $v(t) = v(bb) = v(rb)$
Predicts $\rho(t, X) = \rho(bb, X) = \rho(rb, X) = 1/3$

Intuition $\rho(t, X) = 1/2$ and $\rho(rb, X) = \rho(bb, X) = 1/4$

Correcting **i.i.d.**

Solution 1: Return to random preference
ex.) Assign $\mu = 0.25$ to all of the following orderings

$$t \succ bb \succ rb$$

$$t \succ rb \succ bb$$

$$bb \succ rb \succ t$$

$$rb \succ bb \succ t$$

Then we have $\rho(t, X) = 0.5$ $\rho(rb, X) = \rho(bb, X) = 0.25$.

Correcting **i.i.d.**

Solution 1: Return to random preference
ex.) Assign $\mu = 0.25$ to all of the following orderings

$$t \succ bb \succ rb$$
$$t \succ rb \succ bb$$
$$bb \succ rb \succ t$$
$$rb \succ bb \succ t$$

Then we have $\rho(t, X) = 0.5$ $\rho(rb, X) = \rho(bb, X) = 0.25$.

Solution 2: Non-identical ϵ
ex.) $\epsilon(t) \in \{-1, 1\}$ and $\epsilon(rb), \epsilon(bb) \in \{-10, 10\}$ each with uniform probability

Correcting **i.i.d.**

Solution 1: Return to random preference
ex.) Assign $\mu = 0.25$ to all of the following orderings

$$t \succ bb \succ rb$$
$$t \succ rb \succ bb$$
$$bb \succ rb \succ t$$
$$rb \succ bb \succ t$$

Then we have $\rho(t, X) = 0.5$ $\rho(rb, X) = \rho(bb, X) = 0.25$.

Solution 2: Non-identical ϵ
ex.) $\epsilon(t) \in \{-1, 1\}$ and $\epsilon(rb), \epsilon(bb) \in \{-10, 10\}$ each with uniform probability

Solution 3: Allow correlation on ϵ

1. *i.i.d.* ARU
2. Fechnerian Model
3. Logit Model
   3-1. Logit (or Luce) Model
   3-2. Mixed Logit
   3-3. Nested Logit
   3-4. Perturbed Utility

Psychometric choice metrics

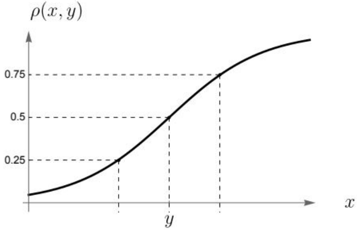

Figure 1.1. An S-shaped psychometric function.

Discrete choice predicts with the step choice rule:

$$\rho(x, y) = \mathbf{1}(x > y) + 0.5 * \mathbf{1}(x = y)$$

Stylized facts from the lab: S-shaped ρ

Causes of randomness: Decision/discrimination errors

Fechnerian Representation

Definition(Fechnerian) ρ has a Fechnerian representation if $\exists v : X \rightarrow \mathbb{R}$ and $\exists F : D \rightarrow [0,1]$ such that

$$\rho(x,y) = F(v(x) - v(y))$$

where $D \equiv \{v(x) - v(y) | x, y \in X\}$ and F: strictly increasing and symmetric at 0.

Other Psychometric Patterns

Examples of non-Fechnerian: Not monotonic in value diff.

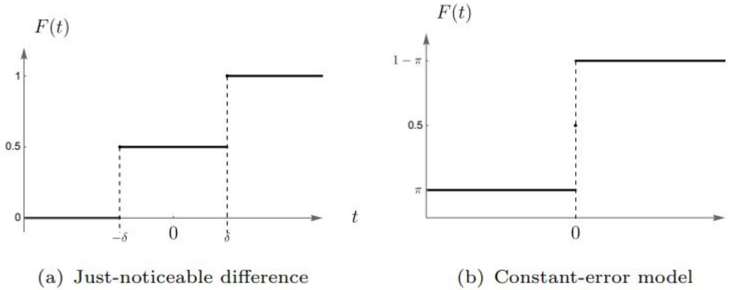

Figure 3.3. Examples of F that is not Fechnerian.

Fechnerian-Axioms

Axiom(Quadraple Condition) For any $x, y, w, z \in X$, $\rho(x, y) \geq \rho(w, z)$ if and only if $\rho(x, w) \geq \rho(y, z)$

Axiom(Richness) If $\rho(x, y) \leq \alpha \leq \rho(z, y)$, then there exists $w \in X$ such that $\rho(w, y) = \alpha$.

Fechnerian-Axioms

Axiom(Quadraple Condition) For any $x, y, w, z \in X$, $\rho(x, y) \geq \rho(w, z)$ if and only if $\rho(x, w) \geq \rho(y, z)$

Axiom(Richness) If $\rho(x, y) \leq \alpha \leq \rho(z, y)$, then there exists $w \in X$ such that $\rho(w, y) = \alpha$.

Relationship with other models

- *i.i.d. ARU*: More general (Equivalent with finite X)
- RU: Less general

Theorem 3.21 (Debreu 1958) Suppose ρ satisfies richness. Then, it has a continuous Fechnerian representation if and only if it satisfies Quadraple Condition. Moreover, if (v_1, F_1) and (v_2, F_2) both represent ρ, then there exists $\alpha > 0$ and β such that

(i) $v_2 = \alpha v_1 + \beta$

(ii) $F_2(\alpha t) = F_1(t)$

Only if part:

$$v(x) - v(y) \geq v(w) - v(z) \Leftrightarrow v(x) - v(w) \geq v(y) - v(z)$$

If part: Use richness for all binary relations

cf.) Non-richness: Replace richness with acyclicity

Acyclicity

Axiom (Acyclicity) For all n, sequences $\{x_i\}_{i=1}^{n}$ and $\{y_i\}_{i=1}^{n}$, and bijections $f, g : \{1, \cdots, n\} \to \{1, \cdots, n\}$,

$$\rho(x_k, y_k) \geq \rho(x_{f(k)}, y_{f(k)}) \text{ for } 1 \leq k < n \Rightarrow \rho(x_n, y_n) \leq \rho(x_{f(n)}, y_{f(n)})$$

Violation of acyclicity: Cycle of $x_1, \cdots, x_{\tilde{n}}$ and $y_1, \cdots, y_{\tilde{n}}$ such that

$$\rho(x_k, x_{k+1}) \geq \rho(y_k, y_{k+1}), \forall 1 \leq k < \tilde{n} \text{ and}$$
$$\rho(x_{\tilde{n}}, x_1) \geq \rho(y_{\tilde{n}}, y_1)$$

Outline

Logit Model

Definition (Logit Choice) If ρ is a logit, then there exists $v : X \to \mathbb{R}$ such that

$$\rho(x, A) = \frac{e^{v(x)}}{\sum_{x' \in X} e^{v(x')}}$$

Definition (Luce Representation) ρ has a Luce representation if there exists $w : X \to \mathbb{R}_{++}$ such that

$$\rho(x, A) = \frac{w(x)}{\sum_{x' \in X} w(x')}$$

Two models are equivalent if $v(x) = \log(w(x))$

Logit-Axioms

Properties allowing inferences on X from $A \subset X$

Logit-Axioms

Properties allowing inferences on X from $A \subset X$

Axiom (Luce IIA) For all $x, y \in A \cap B$, ρ satisfies IIA if

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

Allows predictions from **binary** menu choices

Logit-Axioms

Properties allowing inferences on X from $A \subset X$

Axiom (Luce IIA) For all $x, y \in A \cap B$, ρ satisfies IIA if

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

Allows predictions from **binary** menu choices

Axiom (Luce's Choice Axiom) Let $\rho(A, B) \equiv \sum_{x \in A} \rho(x, B)$. Then, $\forall x \in A \subseteq B$,

$$\rho(x, B) = \rho(x, A)\rho(A, B)$$

where $\rho(A, B) \equiv \sum_{x \in A} \rho(x, B)$

Properties of Logit

Theorem 3.5 The followings are equivalent:

(i) ρ has a logit representation

(ii) ρ has a Luce representation with $w(x) = exp\{v(x)\}$

(iii) ρ satisfies Luce's IIA and positivity

(iv) ρ satisfies Luce's choice axiom and positivity

Mixed Logit

Idea: Bus example solution 1 $\Rightarrow$ Allow randomness over $v(x)$

Definition (Mixed Logit) ρ has a mixed logit representation if

$$\rho(x, A) = \int \frac{e^{v(x)}}{\sum_{x' \in A} e^{v(x')}} d\mu(v)$$

ρ under mixed logit: Expected choice probabliity over $v(x)$

If ρ has a mixed logit representation, it is also RU representable

Caveat) Mixed logit $\neq$ Logit

Mixed Logit-Generalizations

Proposition 3.15 If $\rho \sim RU$, then there is a sequence $\rho^n \sim$ *mixed logit* such that $\rho^n(x, A) \to \rho(x, A)$ for all $x \in X$ and $A \subseteq \mathcal{A}$.

Mixed Logit-Generalizations

Proposition 3.15 If $\rho \sim RU$, then there is a sequence $\rho^n \sim$ *mixed logit* such that $\rho^n(x, A) \to \rho(x, A)$ for all $x \in X$ and $A \subseteq \mathcal{A}$.

Step 1 For any $\succsim \in \mathcal{P}$, assign utility function v

Step 2 Define $\rho^n_v = \dfrac{e^{nv(x)}}{\sum_{x' \in A} e^{nv(x')}}$ and $\rho^n = \sum_v \rho^n_v \mu(v)$

Step 3 Let $\rho^*_v \equiv \mathbf{1}(x = argmax_{x' \in A}(v(x')) = \lim_{n \to \infty} \rho^n_v$. Thus,

$$\rho(x, A) = \sum_v \rho^*_v \mu(v) = \sum_v \lim_{n \to \infty} \rho^n_v \mu(v) = \lim_{n \to \infty} \rho^n(x, A)$$

Nested Logit

Solution 3 of the bus example

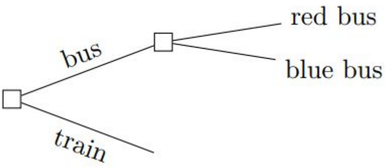

Figure 3.1. A nested decision problem.

Nested Logit

Solution 3 of the bus example

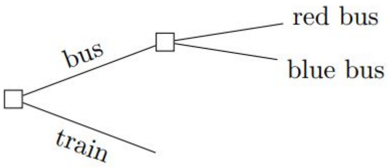

Figure 3.1. A nested decision problem.

Nested logit: Sequential application of logit model

- $\rho(rb, X) = \rho_1(bus, X)\rho_2(rb, bus)$
- ρ_1, ρ_2: Choice probability of each characteristic

Nested Logit-Prediction

Choice probability: BWI of each characteristic choices

Step 1 Logit model between lower-level characteristics
Compute inclusive value of lower-level characteristics

$$V(A) = \mathbb{E}[\max_{x \in A} v(x) + \epsilon(x)] = \log(\sum_{x \in A} \exp v(x))$$

Step 2 Choice among higher-level characteristics
Use the inclusive value for non-random value choice

Definition (Nested Logit) Suppose $x \in B_i \in \mathcal{B}$. ρ is represented by a nested logit model if $\exists v$ such that

$$\rho(x, A) = \frac{e^{V(B_i)}}{\sum_{B_j \in \mathcal{B}} e^{V(B_j)}} \times \frac{e^{v(x)}}{\sum_{x' \in B_i} e^{v(x')}}$$

Model Choice

Pros: Allows structural correlation within the nest

Model Choice

Pros: Allows structural correlation within the nest

Cons: Requires (subjective) judgment on the structures

- Empirical: Gul, Natenzon, and Pesendorfer's (2014)
- Axiomatic: Fudenberg, Iijima, and Strzalecki (2014), Kovach and Tserenjigmid (2022)

Perturbed Utility

Definition (Perturbed Utility) ρ has a PU representation if for each $A \in \mathcal{A}$,

$$\rho(\cdot, A) = \underset{p \in \Delta(A)}{argmax} \sum_{x \in A} v(x)p(x) - c(p)$$

where $c : \Delta(X) \to (-\infty, \infty]$ is a cost function of implementing p.

Interpretation of $c(p)$

(i) Individual makes costly but controllable mistakes (trembling hands)

(ii) Individual probabilitically attends on alternative x (RI)

Relationship between PU and Logit

Literature often assumes an entropy (attention) cost function:

$$c(p) \equiv -\sum_{x \in X} p(x) log(x)$$

Proposition 3.36 The followings are equivalent:

(i) ρ has a logit representation with utility v

(ii) ρ has a PU representation with utility v and Entropy cost

Not only that, ρ has logit representation if $\rho \sim$ *i.i.d.* ARU with T1EV errors

$$\Phi(x) = \exp\{\exp\{-\epsilon(x)\}\}$$

Summary

1. *i.i.d.* ARU model
    - Generalization from pairwise comparison
    - Unique to affine transformation
    - Fails to consider menu dependence

2. Logit model
    - Satisfies IIA and Luce Axiom
    - Mixed/Nested Logit: Allows correlation
    - Rationalizable with entropy attention cost

Appendix

Debreu, G. (1958): "Stochastic Choice and Cardinal Utility", *ECMA*, 26(3), 440¯D444. 45, 176, 177

Fudenberg, D., P. Strack, and T. Strzalecki (2015): "Stochastic Choice and Optimal Sequential Sampling," arXiv preprint arXiv:1505.03342.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *ECMA*, 82(5), 1873–1912

Kovach, M., and G. Tserenjigmid (2022a): "Behavioral foundations of nested stochastic choice and nested logit," *JPE*, 130(9), 2411–2461.

Luce, D. (1959): Individual choice behavior. John Wiley