“Stochastic Choice Theory” by Tomasz Strzalecki
Chapter 1 & 2

Changkuk Im

Department of Economics
Ohio State University

Theory/Experimental Reading Group
May 21, 2024

Chapter 1

Random Utility

Preliminaries

- X: set of all possible alternatives
  - Typically, alternatives are denoted by $x, y, z \in X$
- $\mathcal{A}$: collection of all nonempty and finite subsets of X
  - Typically, menus are denoted by $A, B, C \in \mathcal{A}$
- A single-valued **choice function** is a mapping

$$\chi : \mathcal{A} \to X$$

such that $\chi(A) \in A$

Preliminaries

- X: set of all possible alternatives
  - Typically, alternatives are denoted by $x, y, z \in X$
- $\mathcal{A}$: collection of all nonempty and finite subsets of X
  - Typically, menus are denoted by $A, B, C \in \mathcal{A}$
- A single-valued **choice function** is a mapping

$$\chi : \mathcal{A} \to X$$

  such that $\chi(A) \in A$
  - E.g., $\chi(\{x, y\}) = x$

Preliminaries

- $\rho(x, A)$: frequency with which x from A was observed
- $\Delta(Z)$: set of probability distributions over (a finite set) Z
- A **stochastic choice function (s.c.f.)** is a mapping

$$\rho : \mathcal{A} \to \Delta(X)$$

such that $\sum_{x \in A} \rho(x, A) = 1$ for all $A \in \mathcal{A}$

Preliminaries

- $\rho(x, A)$: frequency with which x from A was observed
- $\Delta(Z)$: set of probability distributions over (a finite set) Z
- A **stochastic choice function (s.c.f.)** is a mapping

$$\rho : \mathcal{A} \to \Delta(X)$$

such that $\sum_{x \in A} \rho(x, A) = 1$ for all $A \in \mathcal{A}$

- ▶ How to interpret $\rho(x, A)$?
- ✓ Individual randomness: fraction of times the agent chose x from A
- ✓ Heterogeneity of preferences: fraction of the populations choosing x from A
- ∗ The classical approach treats that alternatives are *indifferent*

Examples

All menus are observable

Let $X = \{x, y, z\}$ be given.

Then

$$\rho(x, \{x\}) = 1,\ \rho(y, \{y\}) = 1,\ \rho(z, \{z\}) = 1$$
$$\rho(x, \{x, y\}) = \frac{3}{10},\ \rho(y, \{x, y\}) = \frac{7}{10}$$
$$\rho(x, \{x, z\}) = \frac{1}{10},\ \rho(z, \{x, z\}) = \frac{9}{10}$$
$$\rho(y, \{y, z\}) = \frac{7}{10},\ \rho(z, \{y, z\}) = \frac{3}{10}$$
$$\rho(x, \{x, y, z\}) = \frac{1}{10},\ \rho(y, \{x, y, z\}) = \frac{8}{10},\ \rho(z, \{x, y, z\}) = \frac{1}{10}$$

Examples

Only binary menus are observable

Let $X = \{x, y, z\}$ be given.

Then

$$\rho(x, \{x, y\}) = \frac{1}{10},\ \rho(y, \{x, y\}) = \frac{9}{10}$$
$$\rho(x, \{x, z\}) = \frac{3}{10},\ \rho(z, \{x, z\}) = \frac{7}{10}$$
$$\rho(y, \{y, z\}) = \frac{3}{10},\ \rho(z, \{y, z\}) = \frac{7}{10}$$

Models

1. Random Utility
2. Learning
3. Random Consideration
4. Trembling Hands
5. Deliberate Randomization

Models

✓ Random Utility

2. Learning
3. Random Consideration
4. Trembling Hands
5. Deliberate Randomization

Random Utility

There are *three* equivalent ways to formulate the model:

1. Probability distribution over preferences
2. Probability distribution over utility functions
3. Random utility functions

Distribution over Preferences

- $\mathcal{P}$: set of all *strict* preferences over a *finite* set X
  - E.g., $X = \{x, y, z\}$
    $$\begin{aligned}\mathcal{P} = \{& x \succsim y \succsim z, \quad y \succsim x \succsim z, \quad z \succsim x \succsim y, \\ & x \succsim z \succsim y, \quad y \succsim z \succsim x, \quad z \succsim y \succsim x\ \} \\ = \{& xyz, xzy, yxz, yzx, yxz, zyx\}\end{aligned}$$
- $\mu \in \Delta(\mathcal{P})$: probability distribution over strict preferences
  - E.g., $X = \{x, y, z\}$
    $\mu(xyz) = \frac{1}{10}$, $\mu(yxz) = \frac{1}{10}$, $\mu(zxy) = \frac{1}{10}$
    $\mu(xzy) = \frac{1}{10}$, $\mu(yzx) = \frac{1}{10}$, $\mu(zyx) = \frac{5}{10}$

Distribution over Preferences

- For any $A \in \mathcal{A}$ and $x \in A$, let

$$N(x, A) := \{\succsim \in \mathcal{P} : x \succsim y \text{ for all } y \in A\}$$

be the set of *preferences* that rationalizes the choice of x from A

Distribution over Preferences

- For any $A \in \mathcal{A}$ and $x \in A$, let

$$N(x, A) := \{\succsim \in \mathcal{P} : x \succsim y \text{ for all } y \in A\}$$

be the set of *preferences* that rationalizes the choice of x from A

  - E.g., $X = \{x, y, z\}$
    $N(x, \{x, y, z\})$
    $N(x, \{x, y\})$

Distribution over Preferences

- For any $A \in \mathcal{A}$ and $x \in A$, let

$$N(x, A) := \{\succsim \in \mathcal{P} : x \succsim y \text{ for all } y \in A\}$$

be the set of *preferences* that rationalizes the choice of x from A

  - E.g., $X = \{x, y, z\}$
    $N(x, \{x, y, z\}) = \{xyz, xzy\}$
    $N(x, \{x, y\})$

Distribution over Preferences

- For any $A \in \mathcal{A}$ and $x \in A$, let

$$N(x, A) := \{\succsim \in \mathcal{P} : x \succsim y \text{ for all } y \in A\}$$

be the set of *preferences* that rationalizes the choice of x from A

  - E.g., $X = \{x, y, z\}$
    $N(x, \{x, y, z\}) = \{xyz, xzy\}$
    $N(x, \{x, y\}) = \{xyz, xzy, zxy\}$

Distribution over Preferences

Definition 1.6

A s.c.f. $\rho : \mathcal{A} \to \Delta(X)$ is represented by a **distribution over preferences** if there exists $\mu \in \Delta(\mathcal{P})$ such that

$$\rho(x, A) = \mu(N(x, A))$$

for all $A \in \mathcal{A}$ and $x \in A$

Distribution over Preferences

Example

Consider $x = \{x, y, z\}$

Suppose that the s.c.f. is $\rho(x, \{x, y\}) = \frac{1}{3}$ and $\rho(y, \{x, y\}) = \frac{2}{3}$

Then

$\mu(xyz) = \mu(xzy) = \mu(zxy) = \frac{1}{9}$

$\mu(yxz) = \mu(yzx) = \mu(zyx) = \frac{2}{9}$

rationalize the s.c.f. since

$$\rho(x, \{x, y\}) = \frac{1}{3} = \mu(xyz) + \mu(xzy) + \mu(zxy)$$

and

$$\rho(y, \{x, y\}) = \frac{2}{3} = \mu(yxz) + \mu(yzx) + \mu(zyx)$$

Distribution over Utilities

- $U : X \to \mathbb{R}$ or $U \in \mathbb{R}^X$: Utility function
- For any $A \in \mathcal{A}$ and $x \in A$, let

$$\begin{aligned} N(x, A) :=& \{U \in \mathbb{R}^X : U(x) \geq U(y) \text{ for all } y \in A\} \\ =& \{U \in \mathbb{R}^X : U(x) = \max_{y \in A} U(y)\} \end{aligned}$$

be the set of *utility function* that rationalizes the choice x from A

Definition 1.7

A s.c.f. $\rho : \mathcal{A} \to \Delta(X)$ is represented by a **distribution over utilities** if there exists $\mu \in \Delta(\mathbb{R}^X)$ such that

$$\rho(x, A) = \mu(N(x, A))$$

for all $A \in \mathcal{A}$ and $x \in A$

Random Utility Functions

- $(\Omega, \mathcal{F}, \mathbb{P})$: Probability space
  - $\mathcal{F}$ is a σ-algebra
  - $\mathbb{P}$ is a probability measure

- $\tilde{U} : \Omega \to \mathbb{R}^X$: Random utility function

- For any $A \in \mathcal{A}$ and $x \in A$, let

$$\begin{aligned} N(x, A) :=& \{\omega \in \Omega : \tilde{U}_\omega(x) \geq \tilde{U}_\omega(y) \text{ for all } y \in A\} \\ =& \{\omega \in \Omega : \tilde{U}_\omega(x) = \max_{y \in A} \tilde{U}_\omega(y)\} \end{aligned}$$

  be the *event* that rationalizes the choice x from A

Random Utility Functions

Definition 1.8

A s.c.f. $\rho : \mathcal{A} \to \Delta(X)$ has a **random utility** representation if there exists a random variable $\tilde{U} : \Omega \to \mathbb{R}^X$ such that

$$\rho(x, A) = \mathbb{P}(N(x, A))$$

for all $A \in \mathcal{A}$ and $x \in A$

Equivalent Result

Proposition 1.9

The following are equivalent for a *finite* X:

1. ρ is represented by a distribution over preferences;
2. ρ is represented by a distribution over utilities;
3. ρ has a random utility representation

Additive Random Utility

- $v : X \to \mathbb{R}$: deterministic utility function
  - Also called as the "representative utility" or "systematic utility"
- $\tilde{\varepsilon} : \Omega \to \mathbb{R}^X$: random utility shock
  - Private information of the agent
- We write **additive random utility** by

$$\tilde{U}(x) = v(x) + \tilde{\epsilon}(x)$$

  - An equivalent way to write random utility functions
  - In discrete choice econometrics, the main focus is on estimating the function v based on observations of ρ

Additive Random Utility

Definition 1.11

A s.c.f. $\rho : \mathcal{A} \to \Delta(X)$ has a **additive random utility (ARU)** representation if it has a RU representation with

$$\tilde{U}(x) = v(x) + \tilde{\varepsilon}(x),$$

where $v : X \to \mathbb{R}$ is deterministic and the distribution of $\tilde{\varepsilon}$ is smooth

* $\tilde{\varepsilon}$ is smooth if it has a density

Additive Random Utility

Proposition 1.12

If X is *finite*, then $\rho \sim RU$ if and only if $\rho \sim ARU$

Definition 1.13

$\rho : \mathcal{A} \to \Delta(X)$ has a **logit representation** if it has a ARU representation where $\tilde{\varepsilon}(x)$ are i.i.d. across x with the Type 1 Extreme Value distribution, with cdf $G(x) = exp(-exp(-\varepsilon))$

* Details are in Chapter 3

Chapter 2

Basic Properties

Regularity

Axiom 2.1 (Regularity)

If $x \in A \subseteq B$, then $\rho(x, B) \leq \rho(x, A)$

- When we add new alternatives to a menu (i.e., from A to B), the choice probability of existing alternatives should go down
  - E.g., $\rho(x, \{x, y, z\}) = \frac{3}{10} < \frac{7}{10} = \rho(x, \{x, y\})$
  - Stochastic analogue of Sen's α
- Relationship with RU
  - Testable condition of RU
  - Characterization of RU when $|X| = 3$

Relationship Between Regularity and RU

Testable condition

Proposition 2.2 (Block and Marschak, 1960)

If ρ has a random utility representation, then is satisfies Regularity

- E.g., $\rho(x, \{x, y, z\}) = \frac{5}{10} > \frac{3}{10} = \rho(x, \{x, y\})$
  - ρ is NOT rationalized by RU
- Examples of violations
  1. Choice overloads
  2. Asymmetric dominance effect
  3. Compromise effect

Examples of Regularity Violations

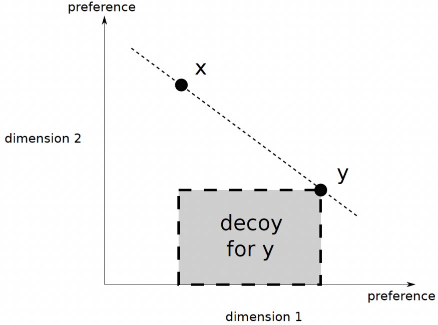
1. Choice overload

Tasting booth in two supermarkets [Iyengar and Lepper, 2000]

- Consumers could (i) taste any number of jams and (ii) buy any variety of jam
  - Supermarket 1: 6 varieties $\Longrightarrow$ 30% purchased
  - Supermarket 2: 24 varieties $\Longrightarrow$ 3% purchased
- ρ(not buying) increased as the menu expanded

Examples of Regularity Violations

2. Asymmetric dominance effect



Hypothetical choices with and without a decoy [Huber et al., 1982]

- Cars, Restaurants, Beers, Lotteries, Films, and TV sets
  - Two attributes (e.g., quality and price)
- $\rho(y, \{x, y, z\})$ increased by 9.2% compared to $\rho(y, \{x, y\})$

Examples of Regularity Violations

2. Asymmetric dominance effect

I. Sample Choice Problem

Below you will find three brands of beer. You know only the price per sixpack and the average quality ratings made by subjects in a blind taste test. Given that you had to choose one brand to buy on this information alone, which one would it be?

Brand	Price/Sixpack	Average Quality Rating (100 = Best; 0 = Worst)
I	$1.80	50
II	$2.60	70
III	$3.00	70

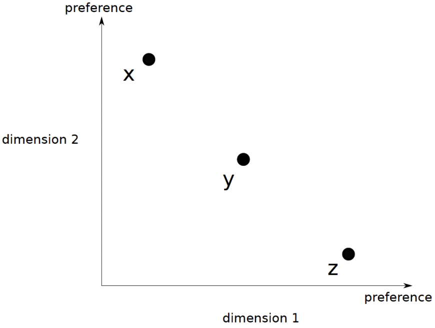
I would prefer Brand—(Check one only)

I ________________ II ________________ III ________________

- Appendix from Huber et al. [1982]

Examples of Regularity Violations

3. Compromise effect



Hypothetical (two-attribute alternative) choices [Simonson, 1989]

- TV, Calculator battery*, Apartment*, Calculator**, Mouthwash**
  - *: One alternative was *unavailable* to choose
  - **: Compromise vs *Extreme*
- $\rho(y, \{x, y, z\})$ increased by 17.5% compared to $\rho(y, \{x, y\})$

Relationship Between Regularity and RU

Characterization when $|X| = 3$

Proposition 2.2 (Block and Marschak, 1960)

If ρ has a random utility representation, then is satisfies Regularity

Proposition 2.3 (Block and Marschak, 1960)

Suppose that $|X| = 3$. If ρ satisfies Regularity, then $\rho \sim RU$

- Unique identification
  - E.g., when $|X| = 3$, $\mu(xyz) = \rho(y, \{y, z\}) - \rho(y, \{x, y, z\})$

When $|X| = 4$

Axiom 2.6 (Supermodularity)

If $x \in A \cap B$, then

$$\rho(x, A) + \rho(x, B) \leq \rho(x, A \cup B) + \rho(x, A \cap B)$$

- The additional impact on $\rho(x, \cdot)$ of adding alternatives to the menu is decreasing in the size of the menu
  - E.g., let $A = \{x, y\}$ and $B = \{x, z\}$
    Then $\rho(x, \{x, y\}) - \rho(x, \{x, y, z\}) \leq \rho(x, \{x\}) - \rho(x, \{x, z\})$

Proposition 2.7 (Block and Marschak)

Suppose that $|X| = 4$. Then ρ satisfies Regularity and Supermodularity if and only if $\rho \sim RU$

BM Polynomials

Axiom 2.8 (Block and Marschak)

For all $x \in A$,

$$q(x, A) := \sum_{B \supseteq A} (-1)^{|B \setminus A|} \rho(x, B) \geq 0$$

- Related to Regularity, Supermodularity, ...
  - $X = \{x, y, z\}$
    $q(x, \{x, y\}) \geq 0$: Regularity
    $q(x, \{x\}) \geq 0$: Supermodularity
  - $X = \{x, y, z, w\}$
    $q(x, \{x, y, z\}) \geq 0$: Regularity
    $q(x, \{x, y\}) \geq 0$: Supermodularity

Characterization of RU

Theorem 2.12

The following conditions are equivalent for ρ on a finite set X:

1. $\rho \sim RU$
2. ρ satisfies the BM axiom
3. ρ satisfies coherency
4. ρ satisfies Axiom of Revealed Stochastic Preference

- (1) $\Longrightarrow$ (2): $q(x, A)$ is the probability of the event that (i) x is the best in A but (ii) everything outside of A is better than x
  - E.g., $X = \{x, y, z, w\}$ and $A = \{x, y\}$

    $$q(x, A) = \mu(\underbrace{z \succ w}_{A^c} \succ \underbrace{x \succ y}_{A}) + \mu(\underbrace{w \succ z}_{A^c} \succ \underbrace{x \succ y}_{A})$$

Uniqueness of RU

Proposition 2.13 (Block and Marshak, 1960)

Suppose that $|X| \leq 3$. If μ is a distribution over preferences that represents ρ, then μ is unique.

- May not be uniquely identifiable when $|X| > 3$
  - E.g., $X = \{x, y, z, w\}$
    $\mu_1(y \succ x \succ w \succ z) = \mu_1(x \succ y \succ z \succ w) = \frac{1}{2}$
    $\mu_2(y \succ x \succ z \succ w) = \mu_2(x \succ y \succ w \succ z) = \frac{1}{2}$
    Then μ_1 and μ_2 generate the same s.c.f.

- Unique identification w/ more structures
  - Single-crossing property [Apesteguia et al., 2017]
  - Branching-independence [Suleymanov, 2024]

Beyond the Book

Recent topics

1. Statistical test of RU
   - Kitamura and Stoye [2018]
2. When ρ is not rationalized by RU
   - Apesteguia and Ballester [2021]
3. Allowing irrational types
   - Filiz-Ozbay and Masatlioglu [2023]
   - Im and Rehbeck [2022], Caliari and Petri [2024]

Conclusion

- Definitions of RU
  - Distribution over preferences (✓)
  - Distribution over utility functions
  - Random utility functions
- Regularity
  - Testable condition
  - Examples of violations: Choice overload and decoy effects
- BM inequality
  - Characterization of RU
  - Related to Regularity, Supermodularity, ...

References I

Jose Apesteguia and Miguel A Ballester. Separating predicted randomness from residual behavior. *Journal of the European Economic Association*, 19(2):1041–1076, 2021.

Jose Apesteguia, Miguel A Ballester, and Jay Lu. Single-crossing random utility models. *Econometrica*, 85(2):661–674, 2017.

Daniele Caliari and Henrik Petri. Irrational random utility models. *arXiv preprint arXiv:2403.10208*, 2024.

Emel Filiz-Ozbay and Yusufcan Masatlioglu. Progressive random choice. *Journal of Political Economy*, 131(3):716–750, 2023.

Joel Huber, John W Payne, and Christopher Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of consumer research*, 9(1):90–98, 1982.

Changkuk Im and John Rehbeck. Non-rationalizable individuals and stochastic rationalizability. *Economics Letters*, 219:110786, 2022.

References II

Sheena S Iyengar and Mark R Lepper. When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology*, 79(6):995, 2000.

Yuichi Kitamura and Jörg Stoye. Nonparametric analysis of random utility models. *Econometrica*, 86(6):1883–1909, 2018.

Itamar Simonson. Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research*, 16(2):158–174, 1989.

Elchin Suleymanov. Branching-independent random utility model. 2024.