Problem Set 03: Choice Functions, WARP and Stochastic Choice

Goes with Lectures 2 and 3. You test whether choices from menus can be explained by a utility function, adapt WARP to indifference, connect it with Sen's conditions α and β, and then move to choice probabilities: random utility, regularity, the logit model, and how to infer a choice correspondence from choice frequencies.

Notation (as in the lectures). X is a finite set of alternatives and $\mathcal{A}$ the set of nonempty menus $A \subseteq X$. A choice function C picks one option $C(A) \in A$; a choice correspondence picks a nonempty set $C(A) \subseteq A$. A stochastic choice function gives probabilities $p(x \mid A)$ with $\sum_{x \in A} p(x \mid A) = 1$; for pairs, $p(x, y) = p(x \mid \{x, y\})$. Logit with utility u and noise $\lambda > 0$: $p(x \mid A) = e^{u(x)/\lambda} / \sum_{y \in A} e^{u(y)/\lambda}$.

Problem 1 Can these choices be explained? Warm-up

Let $X = \{a, b, c, d\}$. Each dataset below lists some menus and the option chosen from each; the other menus were never offered. A utility u *rationalizes* a dataset if, on every observed menu A, the chosen option has a strictly higher utility than every other option in A.

Dataset	Observed choices
D_1	$C(\{a,b\}) = a, \quad C(\{b,c\}) = b, \quad C(\{a,c\}) = c$
D_2	$C(\{a,b\}) = a, \quad C(\{a,b,c\}) = b$
D_3	$C(\{a,b,c,d\}) = b, \quad C(\{a,c\}) = c, \quad C(\{a,d\}) = a, \quad C(\{c,d\}) = c$
D_4	$C(\{a,b\}) = a, \quad C(\{c,d\}) = c$

(a) For each dataset, give a utility function that rationalizes it, or prove that none exists.

(b) For D_3: what does a rationalizing u predict for the menus $\{a, b\}$ and $\{a, c, d\}$? Is the prediction the same for *every* rationalizing u? Same question for D_4 and the menu $\{a, c\}$.

(c) Say that x is *directly revealed preferred* to y, written $x \; R \; y$, if $x \neq y$ and x was chosen from some observed menu that contains y. Prove that a dataset can be rationalized iff R has no cycle $x_1 \; R \; x_2 \; R \cdots R \; x_k \; R \; x_1$. Proceed in two steps.

(i) Show that if u rationalizes the data, then $x \; R \; y$ implies $u(x) > u(y)$. Conclude that there is no cycle.

(ii) Suppose there is no cycle. Show that some $x \in X$ is revealed preferred to no other element. Give it utility 0, remove it, and repeat on the remaining set (the next element removed gets utility 1, and so on). Show that the resulting u rationalizes the data.

(d) Show that D_1 satisfies WARP on the observed menus. Why does this not contradict the lecture's theorem ("WARP $\iff$ rationalizable")?

Problem 2 Choice correspondences: adapting WARP to indifference Core

Let $X = \{x, y, z\}$, and consider two choice correspondences (on singletons, $C(\{w\}) = \{w\}$):

Menu	$\{x,y\}$	$\{x,z\}$	$\{y,z\}$	$\{x,y,z\}$
C_1	$\{x,y\}$	$\{x\}$	$\{y\}$	$\{x,y\}$
C_2	$\{x,y\}$	$\{x\}$	$\{y,z\}$	$\{x\}$

(a) Show that C_1 is generated by a utility: $C_1(A) = \{w \in A : u(w) \geq u(v)\ \forall v \in A\}$ for some u. Then show that C_1 violates the "naive" WARP, copied from choice functions: *if $x \neq y$ are in $A \cap B$ and $x \in C(A)$, then $y \notin C(B)$.* What goes wrong with the naive version?

(b) The lecture's WARP for correspondences reads: for all $A, B \in \mathcal{A}$ and $x, y \in A \cap B$, if $x \in C(A)$ and $y \in C(B)$, then $x \in C(B)$. Show that every correspondence generated by a utility u satisfies it.

(c) Does C_2 satisfy WARP? Can it be generated by a utility?

(d) Show that if $C(A)$ has exactly one element for every A, the lecture's WARP for correspondences is equivalent to WARP for choice functions.

(e) *Arrow's theorem (Arrow, 1959), the hard direction.* Suppose C satisfies WARP, and define $x \succsim^\star y$ iff $x \in C(\{x, y\})$. Proceed in three steps.

(i) Show that $\succsim^\star$ is complete.

(ii) Show that $C(A) = \{x \in A :\ x \succsim^\star y\ \forall y \in A\}$ for every menu A.

(iii) Show that $\succsim^\star$ is transitive. (Use the menu $\{x, y, z\}$ and (ii).)

Conclude with the lecture's weak-order theorem that C is generated by a utility.

Problem 3 Sen's conditions α and β Core

Sen (1971) splits WARP into two conditions on a choice correspondence C. For all menus $A \subseteq B$:

- **Condition α (contraction):** if $x \in A$ and $x \in C(B)$, then $x \in C(A)$.
- **Condition β (expansion):** if $x, y \in C(A)$, then $x \in C(B) \iff y \in C(B)$.

(a) Sen illustrates them with sports: B is the world, A is Pakistan. Translate α and β into statements about world champions and champions of Pakistan.

(b) Show that WARP (for correspondences, as in Problem 2) implies α and β.

(c) Show that α and β together imply WARP. (Hint: given $x, y \in A \cap B$, look at the menu $\{x, y\}$.)

(d) Show that neither condition implies WARP on its own: check that C_2 from Problem 2 satisfies α but not β, and find a correspondence that satisfies β but not α.

(e) Deduce that for choice *functions* (one option per menu), WARP is equivalent to α alone.

(f) A polite guest never takes the last apple. From a basket with one apple a and one orange o, she takes the orange. From a basket with two apples a, a' and one orange, she takes an apple. Which condition does she violate? Is she irrational?

Problem 4 Random utility and regularity Warm-up

Let $X = \{x, y, z\}$. The six strict rankings of X are written xyz (meaning $x \succ y \succ z$), xzy, yxz, yzx, zxy, zyx. A population of students chooses with the following frequencies:

$p(x \mid X)$	$p(y \mid X)$	$p(z \mid X)$	$p(x, y)$	$p(x, z)$	$p(y, z)$
0.5	0.3	0.2	0.6	0.7	0.6

(a) Find a probability distribution π over the six rankings that generates these data, i.e. show that p is a random utility model (RUM). Is π unique?

(b) Check regularity on these data.

(c) In another experiment, $p(x, y) = 0.5$ and, when a third option d (worse than x in every respect) is added, $p(x \mid \{x, y, d\}) = 0.6$, $p(y \mid \{x, y, d\}) = 0.4$. Show directly, without quoting the lecture's result, that no RUM generates these data.

(d) Suppose p is deterministic: every $p(x \mid A)$ is 0 or 1, and $C(A)$ denotes the option with probability 1. Show that p satisfies regularity iff C satisfies condition α of Problem 3. Deduce that a deterministic p is a RUM iff C satisfies WARP.

(e) Show that a mixture of two RUMs, $t\,p_1 + (1-t)\,p_2$ with $t \in [0, 1]$, is a RUM. A market researcher observes that adding d raised the share of x, as in (c). What can she conclude about the individual consumers?

Problem 5 Logit: what the noise parameter λ does

Core

Robinson chooses between three dishes with $u(a) = 4$, $u(b) = 3$, $u(c) = 0$, by logit with noise $\lambda > 0$.

(a) Compute $p(\cdot \mid \{a, b, c\})$ for $\lambda = 0.5$, 1, 2 and 10.

(b) Show that as $\lambda \to 0$, $p(x \mid A) \to 1$ if x is the unique maximizer of u on A. What is the limit if k options tie for the maximum? Show that as $\lambda \to \infty$, $p(x \mid A) \to 1/|A|$.

(c) Show that the probability of the best option on A decreases in λ and that the probability of the worst option increases in λ. Is the probability of b monotone in λ?

(d) Let $\bar{u}(\lambda) = \sum_{x \in A} p(x \mid A)\, u(x)$ be the expected utility of the chosen dish. Show that

$$\bar{u}'(\lambda) = -\frac{1}{\lambda^2} \sum_{x \in A} p(x \mid A)\big(u(x) - \bar{u}(\lambda)\big)^2 \leq 0.$$

Interpret.

(e) Show that (u, λ) and $(\kappa u + m, \kappa\lambda)$, with $\kappa > 0$ and $m \in \mathbb{R}$, give the same choice probabilities. What can be learnt about u from choice probabilities? Compare with Lecture 2, where utility is only ordinal.

(f) *Estimation.* Another student, who also chooses by logit, has $p(a, b) = 0.8$ and $p(b, c) = 0.95$. Predict $p(a, c)$ and $p(\cdot \mid \{a, b, c\})$.

Problem 6 Why Gumbel errors give logit

Challenge

The lecture claims that logit is a random utility model (McFadden, 1974): $U_x = u(x) + \lambda\varepsilon_x$, where the ε_x are independent with the standard *Gumbel* distribution,

$$F(t) = \mathbb{P}(\varepsilon \leq t) = e^{-e^{-t}}, \qquad t \in \mathbb{R}.$$

You may use that $\mathrm{E}[\varepsilon] = \gamma \approx 0.5772$ (Euler's constant) and $\mathrm{Var}(\varepsilon) = \pi^2/6$.

(a) Show that F is a distribution function (increasing, continuous, with limits 0 and 1) and compute its density f.

(b) *Max-stability.* Let $v_1, \ldots, v_n \in \mathbb{R}$ and $\varepsilon_1, \ldots, \varepsilon_n$ independent standard Gumbel. Show that $\max_i (v_i + \varepsilon_i)$ has the same distribution as $\ln\big(\sum_i e^{v_i}\big) + \varepsilon$, where ε is standard Gumbel.

(c) *Two options.* Show that $\mathbb{P}(v_x + \varepsilon_x > v_y + \varepsilon_y) = \dfrac{e^{v_x}}{e^{v_x} + e^{v_y}}$. (Condition on ε_x, and substitute $s = e^{-t}$ in the integral.)

(d) *Any menu.* Use (b) and (c) to show that $\mathbb{P}(v_x + \varepsilon_x > v_y + \varepsilon_y \ \ \forall y \in A \setminus \{x\}) = e^{v_x} / \sum_{y \in A} e^{v_y}$. Why can ties be ignored?

(e) Deduce the logit formula with noise λ from $U_x = u(x) + \lambda\varepsilon_x$. What is the standard deviation of the noise $\lambda\varepsilon_x$?

(f) *Welfare.* Show that the expected utility of the best option, $\mathrm{E}\big[\max_{x\in A} U_x\big]$, equals $\lambda \ln \sum_{x\in A} e^{u(x)/\lambda} + \lambda\gamma$. Show that adding an option to A raises it, and that it tends to $\max_A u$ as $\lambda \to 0$.

(g) Where in the proof is the *independence* of the ε_x used? Explain with the red bus / blue bus example why this assumption is the source of IIA.

Problem 7 From choice frequencies to a choice correspondence

Core

A choice correspondence $C(A)$ is the set of options the decision maker is happy with. But in each trial we see her choose *one* option, never a set. Balakrishnan et al. (2022) propose to infer C from repeated choices. Fix a level of *selectivity* $\delta \in (0, 1]$ (they call it λ; we keep λ for the logit noise) and let

$$C_\delta(A) = \Big\{ x \in A : \; p(x \mid A) \geq \delta \max_{y \in A} p(y \mid A) \Big\}.$$

An option counts as "chosen" unless some other option is chosen more than $1/\delta$ times as often.

(a) Read the frequencies of Problem 4 as the repeated choices of *one* student. Compute C_δ on all menus with at least two options for $\delta = 1$ and $\delta = \frac{1}{2}$. Also compute the "support" rule $C_0(A) = \{x \in A : p(x \mid A) > 0\}$. Which of the three correspondences satisfy WARP (Problem 2)?

(b) Show that on pairs, $x \in C_\delta(\{x, y\})$ iff $p(x, y) \geq \delta/(1 + \delta)$.

(c) Suppose the student chooses by logit with utility u and noise λ. Show that $C_\delta(A) = \{x \in A : u(x) \geq \max_A u - \lambda \ln(1/\delta)\}$. Take $u(a) = 2$, $u(b) = 1$, $u(c) = 0$, $\lambda = 1$, $\delta = \frac{1}{4}$: compute C_δ on all menus and check WARP. Which problem of Problem Set 2 does this remind you of?

(d) In practice we only see n choices. Assume independent trials, and take $\delta = \frac{1}{2}$ and a significance level of 5%. We drop x from $C(\{x, y\})$ only if we can reject $H_0 : \; p(x, y) \geq \frac{1}{3}$, i.e. only if $\mathbb{P}(K \leq k) \leq 0.05$, where k is the number of times x was chosen and $K \sim \mathrm{Bin}(n, \frac{1}{3})$. Decide for: $k = 3$ out of $n = 20$; $k = 2$ out of $n = 20$; $k = 0$ out of $n = 5$.

(e) Tversky (1969) found cycles in choice frequencies: $p(x, y) > \frac{1}{2}$, $p(y, z) > \frac{1}{2}$, $p(z, x) > \frac{1}{2}$. Using (b), explain why such a cycle need not give a cycle in the inferred correspondence C_δ.

References

Arrow, K. J. (1959). Rational choice functions and orderings. *Economica*, 26(102):121–127.

Balakrishnan, N., Ok, E. A., and Ortoleva, P. (2022). Inference of choice correspondences. Working paper, version of December 7, 2022.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

Sen, A. K. (1971). Choice functions and revealed preference. *Review of Economic Studies*, 38(3):307–317.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1):31–48.