Lecture Note on the Theory of Random Utility Models

Kota Saito

August 15, 2019

1 Introduction

The study of stochastic choice is appealing in two ways. First, stochastic choice data are exactly the type of data we observe in empirical analysis. Second, the theory of stochastic choice contains interesting mathematical results that are distinct from those in deterministic choice theory.

However, it can be difficult for a student to gain a unified understanding of the literature on stochastic choice. This difficulty arises from the fact that the literature has developed independently across three different disciplines: psychology, decision theory, and mathematics. In fact, the axiomatization of random utility models was first provided by Falmagne (1978) in mathematical psychology. Without knowing the result, Barbera and Pattanaik (1986) obtained the same axiomatization in economics. Independently, McFadden and Richter (1990) proposed an alternative axiomatization. (They obtained the result in 1970s but the result is published in 1990).

In this note, we will review the classical results achieved by Block and Marschak (1960), Falmagne (1978), and Mcfadden and Richter (1990). Although these results have been regarded as independent of each other, I provide a new unified geometric way to understand these classical results.

2 Random choice function

Let X be a finite set. X is the set of outcomes. Let $\mathcal{D} \equiv 2^X \setminus \{\emptyset\}$.

Definition 1. *A function $\rho : \mathcal{D}\times X \to [0,1]$ is called random choice function if $\sum_{x\in D}\rho(D,x) = 1$ and $\rho(D,x) = 0$ for any $x \notin D$. The set of random choice functions is denoted by $\mathcal{P}$.*

$\rho(D,x)$ is the probability that an element x is chosen from a choice set D. Let $T \equiv |\mathcal{D}| \times |X|$. Note that $|\mathcal{D}| = 2^{|X|} - 1$. A random choice function ρ is an element of $\mathbf{R}^T$.

3 Random utility model

Let V be the set of all functions $v : X \to \mathbf{R}$. Notice that v is $|X|$ dimensional real vector. Let $\mathcal{V}$ be the Borel algebra of V. Denote the finitely additive probability measure by $\Delta(V)$.

Definition 2. *A random choice function ρ is called random utility function if there exists a probability measure $\mu \in \Delta(\mathcal{V})$ such that for all $(D,x) \in \mathcal{D}\times X$,*

$$\rho(D,x) = \mu(v \in \mathcal{V}|v(x) \geq v(D)).$$

Remark 1. *Suppose that a random choice function ρ is random utility function with μ. Then for any $x, y \in X$ such that $\mu(v \in \mathcal{V}|v(x) = v(y)) = 0$.*

Proof. For any $x, y \in X$

$$\begin{aligned}
1 &= \rho(\{x,y\},x) + \rho(\{x,y\},y) \\
&= \mu(v \in \mathcal{V}|v(x) \geq v(y)) + \mu(v \in \mathcal{V}|v(y) \geq v(x)) \\
&= \mu(v \in \mathcal{V}|v(x) \geq v(y)) + \mu(v \in \mathcal{V}|v(y) > v(x)) + \mu(v \in \mathcal{V}|v(y) = v(x)) \\
&= \mu(v \in \mathcal{V}|v(x) \geq v(y) \text{ or } v(y) > v(x)) + \mu(v \in \mathcal{V}|v(y) = v(x)) \\
&= 1 + \mu(v \in \mathcal{V}|v(y) = v(x)).
\end{aligned}$$

So $\mu(v \in \mathcal{V}|v(y) = v(x)) = 0$. ■

Remark 2. *By the finite additivity of μ and the finiteness of X, the above remark implies that the measure of utilities which allows tie is zero.*

3.1 Random ranking model

Let Π be the set of bijection between $X \to \{1, \ldots, |X|\}$. If $\pi(x) = k$, I interpret that x is $|X| + 1 - k$ th best element of X with respect to π. So if $\pi(x) > \pi(y)$, then x is better than y with respect to π. For all $(D, x) \in \mathcal{D} \times X$ if $\pi(x) > \pi(y)$ for all $y \in D \setminus \{x\}$, I write $\pi(x) \geq \pi(D)$. There are $|X|!$ elements in Π.

I denote the set of probability measures over Π by $\Delta(\Pi)$. Since Π is finite, $\Delta(\Pi) = \left\{(\nu_1, \ldots, \nu_{|\Pi|}) \in \mathbf{R}_+^{|\Pi|} \middle| \sum_{i=1}^{|\Pi|} \nu_i = 1\right\}$.

Definition 3. *A random choice function ρ is called random ranking function if there exists a probability measure $\nu \in \Delta(\Pi)$ such that for all $(D, x) \in \mathcal{D} \times X$*

$$\rho(D, x) = \nu(\pi \in \Pi | \pi(x) \geq \pi(D)).$$

The set of random utility functions is denoted by $\mathcal{P}_r$.

Remark 3. *For any random choice function ρ, ρ is a random ranking function if and only if ρ is a random utility function*

Proof. Suppose that ρ is a random utility function to show that ρ is also a random ranking function. For any $\pi \in \Pi$, define

$$\nu(u) = \mu(v \in \mathcal{V} | v(x) > v(y) \text{ if and only if } \pi(x) > \pi(y)).$$

Then for any $(D, x) \in \mathcal{D} \times X$,

$$\rho(D, x) = \mu(v \in \mathcal{V} | v(x) > v(D)) = \nu(\pi \in \Pi | \pi(x) > \pi(D)).$$

To show the converse suppose that ρ is random ranking function to show that ρ is also a random utility function. For any $v \in \mathcal{V}$,

$$\mu(v) = \begin{cases} \nu(\pi) & \text{if } v = \pi, \\ \mu(v) = 0 & \text{otherwise.} \end{cases}$$

Then for any Borel set B define $\mu(B) = \sum_{v \in B \cap \Pi} \mu(v)$. This is well defined because Π is finite. Then for any $(D, x) \in \mathcal{D} \times X$,

$$\rho(D, x) = \nu(\pi \in \Pi | \pi(x) > \pi(D)) = \mu(v \in \mathcal{V} | v(x) > v(D)).$$

■

4 Axiomatization of random utility model by Block-Marschak polynomials

Axiom 1. *(Regularity) For any $D, E \in \mathcal{D}$ such that $x \in D \subset E$, $\rho(D,x) \geq \rho(E,x)$.*

Remark 4. *(i) A random utility function ρ satisfies Regularity. (ii) For any $D \in \mathcal{D}$ such that $x \in D$ and any $z, z' \in X \setminus D$,*

$$\rho(D,x) - \big[\rho(D \cup z, x) + \rho(D \cup z', x)\big] + \rho(D \cup \{z, z'\}, x) \geq 0$$

Proof. (i) holds because

$$\rho(D,x) = \nu(\pi \in \Pi | \pi(x) \geq \pi(D)) \geq \nu(\pi \in \Pi | \pi(x) \geq \pi(D) \;\&\; \pi(x) > \pi(E \setminus D)) = \rho(E,x).$$

To show (ii) holds let $U_1 = \{\pi \in \Pi | \pi(x) \geq \pi(D)\}$, $U_2 = \{\pi \in \Pi | \pi(x) \geq \pi(D) \;\&\; \pi(x) > \pi(z)\}$, $U_2' = \{\pi \in \Pi | \pi(x) \geq \pi(D) \;\&\; \pi(x) > \pi(z')\}$, and $U_3 = \{\pi \in \Pi | \pi(x) \geq \pi(D) \;\&\; \pi(x) > \pi(z) \;\&\; \pi(x) > \pi(z')\}$. Then, $U_3 = U_2 \cap U_2'$. Moreover, $((U_2 \cup U_2') \setminus (U_2 \cap U_2')) \subset U_1$. So $\nu(U_1) - [\nu(U_2) + \nu(U_2')] + \nu(U_3) = \nu(U_1) - [\nu(U_2) + \nu(U_2')] + \nu(U_2 \cap U_2') \geq 0$. ■

Definition 4. *For any $\rho \in \mathcal{P}$ and $(D,x) \in \mathcal{D} \times X$ such that $x \in D$,*

$$K(\rho, D, x) = \sum_{E : D \subset E} (-1)^{|E \setminus D|} \rho(E, x).$$

The number $K(\rho, D, x)$ is called a Block-Marschak polynomial. Block and Marschak (1960) shows that if ρ is a random utility function then the polynomials are nonnegative. Falmagne (1978) shows that the nonnegativity condition is sufficient.

4.1 Necessity of the nonnegativity of B-S polynomials

The following proof is by Barberá and Pattanaik (1986). The proof is based on the inclusion-exclusion formula.

Proposition 1. *For any real valued function h and r defined on a finite set, if*

$$r(T) = \sum_{U : T \subset U} h(U) \tag{1}$$

then

$$h(S) = \sum_{T : S \subset T} (-1)^{|T \setminus S|} r(T). \tag{2}$$

Proof. By substituting $r(T)$ by (1), the right hand side of (2) is

$$\sum_{T:S\subset T}(-1)^{|T\setminus S|}\sum_{U:T\subset U}h(U).$$

For each U, $h(U)$ appears once for each T such that $S\subset T\subset U$ with sign $(-1)^{|T\setminus S|}$. For each $i\in\{0,\dots,|U\setminus S|\}$, the number of T such that $S\subset T\subset U$ and $|T\setminus S|=i$ is

$$\binom{|U\setminus S|}{i}.$$

Therefore, the coefficient on $h(U)$ is

$$\sum_{i=0}^{|U\setminus S|}(-1)^i\binom{|U\setminus S|}{i}.$$

If $U=S$, the coefficient is $(-1)^0\binom{0}{0}=1$. For any U such that $U\neq S$, the coefficient is 0 from binomial formula.

Remember that the binomial formula is $(x+y)^{|U\setminus S|}=\sum_{i=0}^{|U\setminus S|}\binom{|U\setminus S|}{i}x^iy^{|U\setminus S|-i}$. By substituting $x=-1$ and $y=1$,

$$\sum_{i=0}^{|U\setminus S|}(-1)^i\binom{|U\setminus S|}{i}=(-1+1)^{|U\setminus S|}=0.$$

Therefore, the coefficient is 1 only for $h(S)$ and 0 for any $h(U)$ such that $U\neq S$. Thus, the right hand side of (2) is $h(S)$. ■

Proposition 2. *If ρ is a random utility function represented by $\nu\in\Delta(\Pi)$, then for any (D,x) such that $x\in D$,*

$$K(\rho,D,x)=\nu\{\pi\in\Pi|\pi(D^c)>\pi(x)\geq\pi(D)\}.$$

Proof. Fix a nonempty subset D of X and choose an arbitrary element x from D. Define two functions $h,r:2^X\to\mathbf{R}$

$$\begin{aligned}h(E)&=\nu(\pi|\pi(E^c)>\pi(x)\geq\pi(E)),\\ r(E)&=\rho(E,x).\end{aligned}$$

Apply the inclusion and exclusion formula with the two functions h and r. Notice that

$$\{\pi|\pi(x)\geq\pi(D)\}=\bigcup_{E:D\subset E}\{\pi|\pi(E^c)>\pi(x)\geq\pi(E)\}.\tag{3}$$

Moreover, the sets in the right hand side are disjoint. Thus

$$r(D) = \rho(D,x) = \nu(\pi|\pi(x) \geq \pi(D)) = \sum_{E:D\subset E} \nu(\pi|\pi(E^c) > \pi(x) \geq \pi(E)) = \sum_{E:D\subset E} h(E).$$

where the third equality holds by (3). By the inclusion-exclusion formula,

$$\nu\{\pi \in \Pi|\pi(D^c) > \pi(x) \geq \pi(D)\} \equiv h(D) = \sum_{E:D\subset E} r(E)(-1)^{|E\setminus D|} \equiv \sum_{E:D\subset E} \rho(E,x)(-1)^{|E\setminus D|},$$

where the last term equals to $K(\rho, D, x)$. ■

4.2 Falmagne (1978): Sufficiency of nonnegativity of B-S polynomials

Theorem 1. *(Falmagne (1978)) For any $\rho \in \mathcal{P}$, ρ is a random utility function if and only if $K(\rho, D, x) \geq 0$ for any $D \subset X$ and $x \in D$.*

Lemma 1. *For any subset D of X,*

(i) $K(\rho, D, x) = \rho(D, x) - \sum_{E:E\supsetneq D} K(\rho, E, x)$.

(ii) $\sum_{x\in D} K(\rho, D, x) = \sum_{x\in D^c} K(\rho, D\cup\{x\}, x)$.

Proof. *Step* 1: (i). For each F such that $F \supsetneq D$, we show that $\rho(F,x)$ appears once with the sign $(-1)^{|F\setminus D|}$ in the term $-\sum_{E:E\supsetneq D} K(\rho, E, x)$. Therefore, the right hand side coincides with the definition of $K(\rho, D, x)$.

Fix F such that $F \supsetneq D$. In the term $\sum_{E:E\supsetneq D} K(\rho, E, x)$, $\rho(F,x)$ appears once for each E such that $F \supset E \supsetneq D$ with sign $(-1)^{|F\setminus E|}$. Suppose that $|F \setminus E| = i$, then the number of such E must be the same as $\binom{|F\setminus D|}{i}$. Since $E \neq D$, we have $i < |F \setminus D|$. Thus the coefficient of $\rho(F,x)$ is

$$\sum_{i=0}^{|F\setminus D|-1} \binom{|F\setminus D|}{i}(-1)^i = \sum_{i=0}^{|F\setminus D|} \binom{|F\setminus D|}{i}(-1)^i - (-1)^{|F\setminus D|} = -(-1)^{|F\setminus D|}.$$

Step 2: (ii). The left hand side of (ii) becomes:

$$\begin{aligned}
\sum_{x\in D} K(\rho, D, x) &= \sum_{x\in D}\sum_{E:E\supset D}(-1)^{|E\setminus D|}\rho(E,x) \\
&= \sum_{E:E\supset D}(-1)^{|E\setminus D|}\sum_{x\in D}\rho(E,x) \\
&= \sum_{E:E\supset D}(-1)^{|E\setminus D|}\Big(1 - \sum_{y\notin D}\rho(E,y)\Big) \\
&= \sum_{E:E\supset D}(-1)^{|E\setminus D|} - \sum_{E:E\supset D}(-1)^{|E\setminus D|}\sum_{y\notin D}\rho(E,y) \\
&= -\sum_{E:E\supset D}(-1)^{|E\setminus D|}\sum_{y\notin D}\rho(E,y),
\end{aligned}$$