CALIFORNIA INSTITUTE OF TECHNOLOGY
Division of the Humanities and Social Sciences

Introductory notes on stochastic rationality*

KC Border
Fall 2007
v. 2011.10.05::17.46

1 Stochastic choice and stochastic rationality

In the standard theory of rational choice we start with a set X of alternatives, a family $\mathcal{B}$ of nonempty budgets (subsets of X), and a choice function $c\colon B \mapsto c(B) \subset B$. In one model of **stochastic choice** a single element of B is chosen at random.[1] So instead of a subset of B the "choice" gives us a probability measure on B.

Let $p(C|B)$ denote the probability that the choice from B belongs to C.

Of course we require that $p(B|B) = 1$ for all $B \in \mathcal{B}$. This notation is intentionally reminiscent of the notation for the conditional probability C given B, which it may not be, but I don't think you will get confused here. When C and/or B are itemized, we may simply omit the braces. For example, we may write $p(x|x, y)$ instead of the cumbersome $p\big(\{x\}|\{x, y\}\big)$.

To keep things simple, we shall assume that X is finite.

A special kind of stochastic choice is derived from a stochastic preference. Let $\mathcal{P}$ denote the set of linear preference relations on X. (A preference relation $\succ$ is *linear* if it is total, transitive, irreflexive, and asymmetric.) The set $\mathcal{P}$ is also finite, having $|X|!$ elements. We may when convenient refer to a preference relation by listing its "skeleton." For instance, if $X = \{x, y, z\}$, we may write $z \succ x \succ z$ to refer to the unique linear preference $\succ$ satisfying $z \succ x \succ z$.

A **stochastic preference** is a probability measure on $\mathcal{P}$.[2] A stochastic preference gives rise to a stochastic choice as follows. Given a linear preference $\succ$ on X and budget B, let

$$\succ(B) = \text{ the (unique) } \succ\text{-maximal element of } B.$$

1 Definition *For the purposes of this note, a stochastic choice $(X, \mathcal{B}, p)$ is* ***stochastically rational*** *if there is a stochastic preference π on $\mathcal{P}$ such that for all $B \in \mathcal{B}$,*

$$p(x|B) = \pi\big(\{\succ \in \mathcal{P} : x = \succ(B)\}\big).$$

In this case we say that π rationalizes p.

*These notes expand upon Exercise 1.D.5 in Mas-Colell–Whinston–Green [8, p. 16], and are a simplified (I hope) exposition of some results in McFadden and Richter [10].

[1] There are other models of stochastic choice. We could, for instance, assume that a subset of B is chosen at random. Indeed McFadden and Richter [10] consider an entirely general framework. The case considered here is both natural and simple.

[2] A more general, and more complicated, model would allow indifference.

The next example shows that not every stochastic choice can be rationalized by a stochastic preference. It appears as Exercise 1.D.5 in Mas-Colell–Whinston–Green [8, p. 16].

2 Example Let $X = \{x, y, z\}$, and let $\mathcal{B}$ be the set of all two-element subsets of X. Then the stochastic choice

$$\begin{aligned} p(x|x,y) &= \tfrac{1}{2}, \quad p(y|x,y) = \tfrac{1}{2} \\ p(y|y,z) &= \tfrac{1}{2}, \quad p(z|y,z) = \tfrac{1}{2} \\ p(z|x,z) &= \tfrac{1}{2}, \quad p(x|x,z) = \tfrac{1}{2} \end{aligned} \tag{1.1}$$

is stochastically rational, and is rationalized by π where $\pi(\succ) = 1/6$ for each $\succ \in \mathcal{P}$.

But the stochastic choice

$$\begin{aligned} p(x|x,y) &= \tfrac{3}{4}, \quad p(y|x,y) = \tfrac{1}{4} \\ p(y|y,z) &= \tfrac{3}{4}, \quad p(z|y,z) = \tfrac{1}{4} \\ p(z|x,z) &= \tfrac{3}{4}, \quad p(x|x,z) = \tfrac{1}{4} \end{aligned} \tag{1.2}$$

is not stochastically rational. Perhaps the simplest way to see this is to note that if $z \succ x$, then there are three choices for where y fits in the preference order: $y \succ z \succ x$, $z \succ y \succ x$, or $z \succ x \succ y$. In any event, $y \succ x$ or $z \succ y$ (or both). In other words,

$$\{\succ \in \mathcal{P} : z \succ x\} \subset \{\succ \in \mathcal{P} : z \succ y\} \bigcup \{\succ \in \mathcal{P} : y \succ x\}.$$

But if p is rationalized by some probability π on $\mathcal{P}$, then (1.2) implies that left-hand side of the inclusion has probability $3/4$, while the union on the right-hand side has probability at most $1/4 + 1/4$, a contradiction.

Both (1.1) and (1.2) are special cases of the following stochastic choice.

$$\begin{aligned} p(x|x,y) &= \alpha, \quad p(y|x,y) = 1-\alpha \\ p(y|y,z) &= \alpha, \quad p(z|y,z) = 1-\alpha \\ p(z|x,z) &= \alpha, \quad p(x|x,z) = 1-\alpha \end{aligned} \tag{1.3}$$

So when is this stochastically rational? The answer is, whenever

$$1/3 \leqslant \alpha \leqslant 2/3.$$

First we show that this condition is necessary. The same argument used above, namely

$$\{\succ \in \mathcal{P} : z \succ x\} \subset \{\succ \in \mathcal{P} : z \succ y\} \bigcup \{\succ \in \mathcal{P} : y \succ x\}$$

shows that a necessary condition is that $\alpha \leqslant 2(1-\alpha)$, or $\alpha \leqslant 2/3$. But

$$\{\succ \in \mathcal{P} : y \succ x\} \subset \{\succ \in \mathcal{P} : y \succ z\} \bigcup \{\succ \in \mathcal{P} : z \succ x\}$$

implies that $1 - \alpha \leqslant 2\alpha$, or $1/3 \leqslant \alpha$. To show sufficiency, it is enough to exhibit a π that rationalizes p. So assume $1/3 \leqslant \alpha \leqslant 2/3$, and set

$$\begin{aligned} \pi(x \succ y \succ z) = \pi(y \succ z \succ x) = \pi(z \succ x \succ y) = \alpha - 1/3 \\ \pi(x \succ z \succ y) = \pi(y \succ x \succ z) = \pi(z \succ y \succ x) = 2/3 - \alpha. \end{aligned}$$

Simple arithmetic shows that this π rationalizes p. □

2 The Axiom of Revealed Stochastic Preference

The natural follow-up question is what properties must a random choice p satisfy in order to guarantee that it is rationalized by some random preference π?

We shall prove that stochastically rational choice functions are characterized by the following axiom, which McFadden and Richter [10] dub the Axiom of Revealed Stochastic Preference, or ARSP for short.

3 Axiom of Revealed Stochastic Preference *A stochastic choice* $(X, \mathcal{B}, p)$ *satisfies the* ***ARSP*** *if for every finite sequence* $(C_1, B_1), \dots, (C_n, B_n)$ *(where repetitions are allowed) with each* $B_j \in \mathcal{B}$*, and each* $C_j \subset B_j$*,*

$$\sum_{j=1}^{n} p(C_j|B_j) \leqslant \max_{\succ \in \mathcal{P}} \sum_{j=1}^{n} \mathbf{1}(\succ, C_j, B_j) \tag{2.1}$$

where

$$\mathbf{1}(\succ, C, B) = \begin{cases} 1 & \textit{if } \succ(B) \in C \\ 0 & \textit{otherwise.} \end{cases}$$

Or to put it another way that makes the repetitions explicit, let $\{(C_i, B_i) : i = 1, \dots, N\}$ *be an enumeration of all the distinct pairs* (C, B) *such that* $C \subset B \in \mathcal{B}$*. Then for every tuple* $(k_i)_{i=1}^N$ *of nonnegative integers,*

$$\sum_{i=1}^{N} k_i p(C_i|B_i) \leqslant \max_{\succ \in \mathcal{P}} \sum_{i=1}^{N} k_i \mathbf{1}(\succ, C_i, B_i). \tag{2.1$'$}$$

The right-hand side of (2.1) begs for some comment. What it does is find a single preference relation that would make the most choices from B_j lying in C_j. Clearly if no preference relation would ever choose something in C_j from B_j, then no randomly chosen preference would choose something in C_j from B_j, so a stochastically rational choice would assign probability zero to such an event. The remarkable thing is that an upper bound on the sum of the probabilities is found by in terms of the number of choices a single preference relation would make.

The ARSP looks very different in form from our standard revealed preference axioms, but it is not so alien as it first seems. Let's apply it to the singleton-valued non-stochastic choice case. Let $h \colon \mathcal{B} \to X$ be a singleton-valued choice function. Let us abuse notation slightly and write $x = h(B)$ rather than $\{x\} = h(B)$. Then

$$p(C|B) = \begin{cases} 1 & \text{if } h(B) \in C \\ 0 & \text{otherwise} \end{cases}$$

is the corresponding stochastic $\{0, 1\}$-valued choice function. Let S be the strong direct revealed preference relation, defined, as you may recall, by $x \; S \; y$ if for some budget $B \in \mathcal{B}$, we have $y \in B$, $y \neq x$, and $x = h(B)$. Now assume

$$x_1 \; S \; x_2 \; S \; x_3 \cdots S \; x_n,$$

all the x_j distinct,[3] and let B_j be such that $x_j = h(B_j)$ and $x_{j+1} \in B_j$, $j = 1, \dots, n-1$. We now ask, can we have $x_n \; S \; x_1$? If so, let $x_n = h(B_n)$ where $x_1 \in B_n$. Then

$$\sum_{j=1}^{n} p(x_j|B_j) = n.$$

Now consider any linear order $\succ$. Transitivity and irreflexivity rule out $x_1 \succ \cdots \succ x_n \succ x_1$, so

$$\sum_{j=1}^{n} \mathbf{1}(\succ, x_j, B_j) \leqslant n - 1.$$

Thus $x_n \; S \; x_1$ implies that (2.1) is violated, so ARSP implies that

$$x_1 \; S \; x_2 \; S \; x_3 \cdots S \; x_n \implies \neg \, x_n \; S \; x_1,$$

which is just the Strong Axiom of Revealed Preference!

4 Example Let's apply the ARSP to Example 2.

Consider the sequence

$$\big(\{x\}, \{x, y\}\big), \big(\{y\}, \{y, z\}\big), \big(\{z\}, \{x, z\}\big).$$

For this sequence the left-hand side of (2.1) is

$$\alpha + \alpha + \alpha = 3\alpha$$

Now consider the linear preference $\succ$ given by $x \succ y \succ z$. The corresponding sum for the right-hand side of (2.1) is

$$\mathbf{1}\big(\succ, \{x\}, \{x, y\}\big) + \mathbf{1}\big(\succ, \{y\}, \{y, z\}\big) + \mathbf{1}\big(\succ, \{z\}, \{x, z\}\big) = 1 + 1 + 0 = 2.$$

By symmetry, this is the maximum value for the right-hand sum. Thus in order for (2.1) to hold it is necessary that

$$3\alpha \leqslant 2, \quad \text{or} \quad \alpha \leqslant 2/3.$$

Similarly, for the sequence

$$\big(\{y\}, \{x, y\}\big), \big(\{z\}, \{y, z\}\big), \big(\{x\}, \{x, z\}\big),$$

we see that (2.1) implies

$$3(1 - \alpha) \leqslant 2, \quad \text{or} \quad 1/3 \leqslant \alpha.$$

Now we would like to show that if $1/3 \leqslant \alpha \leqslant 2/3$, then (2.1) holds. It is clear that a pair (C, B) with $C = B$ contributes 1 to each side of (2.1), so we need only consider sequences of pairs of the form $\big(\{a\}, \{a, b\}\big)$. Even so, it's not apparent how to proceed. I'll postpone further discussion until after the proof. □

[3] The only thing the assumption of distinctness rules out is $x_n = x_1$, since if $x_i = x_j$ for $i < j$, we may simply omit $x_i, \dots, x_{j-1}$. We deal with the possibility $x_n = x_1$ in the next sentence.

3 Characterization of stochastic rationality

5 Theorem (McFadden–Richter [10]) *A stochastic choice p is rationalized by a random preference π on $\mathcal{P}$ if and only it satisfies ARSP.*

Proof: (ARSP $\implies$ stochastic rationality): We shall prove this part by contraposition.

Let $\mathcal{I}$ be the set of all distinct pairs (C, B) with $B \in \mathcal{B}$ and $\varnothing \neq C \subset B$. Construct the matrix with rows indexed by $\mathcal{I}$, and columns indexed by $\mathcal{P}$. Now append a row of ones. Then p is rationalized by a random preference if and only if the following system of equations has a nonnegative solution $\pi \in \boldsymbol{R}^{\mathcal{P}}$.

$$\begin{array}{c} \\ (C,B) \\ \\ \hline \\ \end{array} \left[\begin{array}{ccc} & \overset{\succ}{\vdots} & \\ \cdots & \mathbf{1}(\succ, C, B) & \cdots \\ & \vdots & \\ \hline \cdots & 1 & \cdots \end{array}\right] \left[\begin{array}{c} \vdots \\ \pi(\succ) \\ \vdots \end{array}\right] = \left[\begin{array}{c} \vdots \\ p(C|B) \\ \vdots \\ \hline 1 \end{array}\right] \tag{3.1}$$

So to prove the contrapositive, we assume that stochastic rationality fails—that is, (3.1) has no solution. By Farkas' Lemma (see, e.g., Gale [4, Theorem 2.6, p. 44]), the alternative is that there exists a vector $y = \big(\cdots, y(C, B), \cdots; y_0\big) \in \boldsymbol{R}^{\mathcal{I}} \times \boldsymbol{R}$ such that

$$\big[\cdots, y(C,B), \cdots; y_0\big] \left[\begin{array}{ccc} & \vdots & \\ \cdots & \mathbf{1}(\succ, C, B) & \cdots \\ & \vdots & \\ \hline \cdots & 1 & \cdots \end{array}\right] \leqq 0, \qquad \big[\cdots, y(C,B), \cdots; y_0\big] \left[\begin{array}{c} \vdots \\ p(C|B) \\ \vdots \\ \hline 1 \end{array}\right] > 0$$

or writing it out, for each $\succ \in \mathcal{P}$,

$$\sum_{(C,B)\in\mathcal{I}} y(C, B)\mathbf{1}(\succ, C, B) + y_0 \leqslant 0, \tag{3.2}$$

and

$$\sum_{(C,B)\in\mathcal{I}} y(C, B)p(C|B) + y_0 > 0. \tag{3.3}$$

Together (3.2) and (3.3) imply that for every $\succ \in \mathcal{P}$,

$$\sum_{(C,B)\in\mathcal{I}} y(C, B)\mathbf{1}(\succ, C, B) < \sum_{(C,B)\in\mathcal{I}} y(C, B)p(C|B) \tag{3.4}$$

Let

$$\mathcal{I}^+ = \{(C, B) \in \mathcal{I} : y(C, B) \geqslant 0\} \quad \text{and} \quad \mathcal{I}^- = \{(C, B) \in \mathcal{I} : y(C, B) < 0\}.$$

Then (3.4) becomes

$$\begin{aligned} \sum_{(C,B)\in\mathcal{I}^+} y(C, B)\mathbf{1}(\succ, C, B) &- \sum_{(C,B)\in\mathcal{I}^-} |y(C, B)|\mathbf{1}(\succ, C, B) \\ &< \sum_{(C,B)\in\mathcal{I}^+} y(C, B)p(C|B) - \sum_{(C,B)\in\mathcal{I}^-} |y(C, B)|p(C|B). \end{aligned} \tag{3.5}$$

Since this is a finite system of strict inequalities, if there is a solution y, then there is a solution where the coordinates of y are rational numbers. By multiplying by a common denominator, we can find a solution y with integer coordinates. Moreover every coefficient $y(C, B) > 0$ for $(C, B) \in \mathcal{I}^+$ and of course $|y(C, B)| > 0$ for $(C, B) \in \mathcal{I}^-$. Define the nonnegative integers

$$k(C,B) = \begin{cases} y(C,B) & (C,B) \in \mathcal{I}^+ \\ |y(C,B)| & (C,B) \in \mathcal{I}^-. \end{cases}$$

Then we may rewrite (3.5) as

$$\begin{aligned} \sum_{(C,B)\in\mathcal{I}^+} k(C,B)\mathbf{1}(\succ, C, B) - \sum_{(C,B)\in\mathcal{I}^-} k(C,B)\mathbf{1}(\succ, C, B) \\ < \sum_{(C,B)\in\mathcal{I}^+} k(C,B)p(C|B) - \sum_{(C,B)\in\mathcal{I}^-} k(C,B)p(C|B). \end{aligned} \tag{3.6}$$

Now observe that $\mathbf{1}(\succ, C, B) = 1 - \mathbf{1}(\succ, B \setminus C, B)$ and $p(C|B) = 1 - p(B \setminus C|B)$, so (3.6) can be written as

$$\begin{aligned} \sum_{(C,B)\in\mathcal{I}^+} k(C,B)\mathbf{1}(\succ, C, B) + \sum_{(C,B)\in\mathcal{I}^-} k(C,B)\mathbf{1}(\succ, B\setminus C, B) \\ < \sum_{(C,B)\in\mathcal{I}^+} k(C,B)p(C|B) + \sum_{(C,B)\in\mathcal{I}^-} k(C,B)p(B\setminus C|B). \end{aligned} \tag{3.7}$$

Now consider the finite collection consisting of $k(C, B)$ instances of (C, B) for $(C, B) \in \mathcal{I}^+$ and $k(C, B)$ instances of $(B \setminus C, B)$ for $(C, B) \in \mathcal{I}^-$. Then since (3.7) holds for each $\succ \in \mathcal{P}$, we have a violation of (2.1′). This proves that if p is not stochastically rational, then ARSP is violated. That is, if p satisfies ARSP, then it is stochastically rational.

(Stochastic rationality $\Longrightarrow$ ARSP): We could run the argument in reverse to prove the converse, but a direct argument is more instructive. Assume that the probability π on $\mathcal{P}$ stochastically rationalizes the stochastic choice p, and consider the collection of distinct pairs $(C_i, B_i)_{i=1}^N$ with k_i instances of each. For each i, by stochastic rationality,

$$p(C_i|B_i) = \sum_{\succ\in\mathcal{P}} \pi(\succ) \cdot \mathbf{1}(\succ, C_i, B_i)$$

so

$$\begin{aligned} \sum_{i=1}^N k_i p(C_i|B_i) &= \sum_{i=1}^N k_i \sum_{\succ\in\mathcal{P}} \pi(\succ) \cdot \mathbf{1}(\succ, C_i, B_i) \\ &= \sum_{\succ\in\mathcal{P}} \pi(\succ) \left(\sum_{i=1}^N k_i \mathbf{1}(\succ, C_i, B_i) \right) \\ &\leqslant \underbrace{\sum_{\succ\in\mathcal{P}} \pi(\succ)}_{=1} \left(\max_{\succ\in\mathcal{P}} \sum_{i=1}^N k_i \mathbf{1}(\succ, C_i, B_i) \right) \\ &= \max_{\succ\in\mathcal{P}} \sum_{i=1}^N k_i \mathbf{1}(\succ, C_i, B_i), \end{aligned}$$

which is just (2.1′). Thus stochastic rationality implies ARSP. ∎

The proof is a bit disappointing because it does not seem to use at all the intuition from the example. It is possible that the symmetry in the example is atypical, and the reasoning there does not generalize well. I need to think hard about this.

4 Effectivity

The ARSP imposes infinitely many restrictions on a stochastic choice even when X is finite, so you might ask whether it is feasible to verify it. In fact, there is a computational procedure for checking stochastic rationality directly in the finite case—the simplex method of linear programming. The system (3.1) of equations is of the form

$$A\pi = p \tag{4.1}$$

where $p \geqq 0$. Consider the following linear program.

$$\text{minimize } \mathbf{1} \cdot z \quad \text{subject to} \quad A\pi + z = p,\ \pi \geqq 0,\ z \geqq 0. \tag{4.2}$$

The system (4.1) has a nonnegative solution $\pi = \bar{\pi}$ if and only if $\pi = \bar{\pi}$, $z = 0$ is a solution of (4.2). In fact, if (4.1) has no solution, then the solution of the dual program will imply (3.2) and (3.3). The program (4.2) is ideally set up to solve with the simplex algorithm since $\pi = 0$, $z = p$ is an obvious initial feasible point.

5 Related literature

Thurstone [11, 12] and Luce [7] do not frame the problem quite the same way we have, but their work paved the way for random utility models. McFadden [9] and Falmagne [3], address the question of stochastic rationality much as we have defined it. Gul and Pessendorfer [5, 6] take the alternatives themselves to be lotteries. There are many other related papers. McFadden [9] has an extensive, if typo-prone, bibliography and literature survey.

References and related literature

[1] H. D. Block and J. Marschak. 1959. Random orderings and stochastic theories of response. Cowles Foundation Discussion Paper 66, Cowles Foundation, New Haven, Connecticut. cowles.econ.yale.edu/P/cp/p01a/p0147.pdf.

[2] R. Corbin and A. A. J. Marley. 1974. Random utility models with equality: An apparent, but not actual, generalization of random utility models. *Journal of Mathematical Psychology* 11(3):274–293. DOI: 10.1016/0022-2496(74)90023-6.

[3] J. C. Falmagne. 1978. A representation theorem for finite random scale systems. *Journal of Mathematical Psychology* 18(1):52–72. DOI: 10.1016/0022-2496(78)90048-2.

[4] D. Gale. 1989. *Theory of linear economic models*. Chicago: University of Chicago Press. Reprint of the 1960 edition published by McGraw-Hill.

[5] F. Gül and W. Pesendorfer. 2006. Random expected utility. *Econometrica* 74(1):121–146. www.jstor.org/stable/3598925.pdf.

[6] ——— . 2006. Supplement to "Random expected utility". www.econometricsociety.org/ecta/supmat/ECTA4734SUPP.pdf.

[7] R. D. Luce. 1958. A probabilistic theory of utility. *Econometrica* 26:193–224. www.jstor.org/stable/1907587.

[8] A. Mas-Colell, M. D. Whinston, and J. R. Green. 1995. *Microeconomic theory*. Oxford: Oxford University Press.

[9] D. L. McFadden. 2004. Revealed stochastic preference: A synthesis. Working paper, University of California (Berkeley). www.econ.berkeley.edu/wp/mcfadden0204/stochastic.pdf.

[10] D. L. McFadden and M. K. Richter. 1990. Stochastic rationality and revealed preference. In J. S. Chipman, D. L. McFadden, and M. K. Richter, eds., *Preferences, Uncertainty, and Optimality: Essays in Honor of Leonid Hurwicz*, pages 163–186. Boulder, Colorado: Westview Press.

[11] L. L. Thurstone. 1927. A law of comparative judgement. *Psychological Review* 34(4):273–286. DOI: 10.1037/h0070288.

[12] ——— . 1927. Psychophysical analysis. *American Journal of Psychology* 38(3):368–389. www.jstor.org/stable/1415006.