Journal of Mathematical Psychology 84 (2018) 89–99

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

Strict scalability of choice probabilities☆

Matthew Ryan

School of Economics, Auckland University of Technology, New Zealand

HIGHLIGHTS

- Weak substitutability is shown to characterise what we call *strict scalability* of binary choice probabilities.
- Strict scalability lies between the classical notions of simple and monotone scalability.
- Strict scalability ensures that the utility scale represents stochastic preferences.
- A multinomial generalisation of weak substitutability that characterises multinomial strict scalability is defined.

ARTICLE INFO

Article history:
Received 20 February 2018
Received in revised form 18 April 2018
Available online 11 May 2018

ABSTRACT

This paper introduces the concept of *strict scalability*, which lies between the classical notions of *simple scalability* (Krantz, 1964; Tversky and Russo, 1969) and *monotone scalability* (Fishburn, 1973). For binary choices, strict scalability is precisely characterised by the well-known axiom of *weak substitutability* (at least for countable domains). We also introduce a multinomial extension of weak substitutability that characterises strict scalability for multinomial choice.

1. Introduction

This paper explores the axiomatic foundations of classical representations of choice probabilities. Our primary purpose is to fill a natural position in the gap between the well-known concept of *simple scalability* (Krantz, 1964; Tversky & Russo, 1969) and its weaker cousin, *monotone scalability* (Fishburn, 1973). Between these two notions lies a new concept that we call *strict scalability*.

The motivation for introducing this new concept is twofold. First, it repairs a "deficiency" in monotone scalability. The latter property is compatible with a utility scale that fails to represent the decision-maker's stochastic preferences. We say that a is "stochastically preferred" to b if the decision-maker chooses a over b more often than she chooses b over a in a binary choice context. A utility scale "represents" stochastic preferences if it assigns higher utility to a than to b if and only if (iff) a is stochastically preferred to b. It is often natural to restrict attention to utility scales with this property. Strict scalability requires monotone scalability *with respect to a utility scale that represents the decision-maker's stochastic preferences*. This is obviously a more stringent requirement than monotone scalability *simpliciter*. We show that there exist monotone scalable choice probabilities that are not strictly scalable, and also that strict scalability is strictly weaker than simple scalability.

Second, when alternatives are drawn from a countable set, the strictly scalable binary choice probabilities are precisely those that satisfy the familiar property of *strong stochastic transitivity (SST)*, which is equivalent to *weak substitutability* (Davidson & Marschak, 1959). Strict scalability therefore gives a convenient characterisation of the latter class of binary choice probabilities. Given the prominence of SST and weak substitutability in the literature on probabilistic choice, it is useful to have such a characterisation.

When the set from which alternatives are drawn need not be countable, SST is necessary but not quite sufficient for strict scalability of binary choice probabilities. Theorem 14 in Section 2.2 provides a set of necessary and sufficient conditions, analogous to those in Fishburn's (1973) representation theorem for monotone scalability (*ibid*., Theorem A).

Section 3 extends our characterisation of strict scalability from binary to multinomial choice probabilities. This complements the work of Krantz (1964), Smith (1976) and Tversky (1972) on multinomial simple scalability. In order to generalise our representation theorem, we introduce a multinomial generalisation of weak substitutability that we call (not very imaginatively) *multinomial weak substitutability* — see Definition 25.

The Appendix contains definitions of some standard properties of binary choice probabilities, along with well-known results on simple and monotone scalability in the binary choice context.

☆ My thanks to two anonymous referees for their careful reading of the manuscript, their correction of errors in the previous version and their suggestions for improvement in the exposition. Special thanks are owed to Tony Marley, who gave generously of his time to provide detailed and extremely helpful comments on earlier drafts of this paper.

E-mail address: mryan@aut.ac.nz.

https://doi.org/10.1016/j.jmp.2018.04.001
0022-2496/

Figs. 1 and 2 summarise the relationships discussed in the Appendix plus the new results on binary choice established in the present paper. These figures elaborate (parts of) Figure 1 in Fishburn (1973), and follow similar principles in their construction — these principles are explained in Section 2.3.

2. Binary choice probabilities

Given a non-empty set A of alternatives, a *complete binary choice specification (CBCS)* gives the probability with which each alternative is chosen from any binary choice set $\{a, b\} \subseteq A$. The "completeness" qualifier emphasises that choice probabilities are defined for *any* doubleton subset of A.[1] As a formal matter, we also define choice probabilities for "pairs" containing two replicas of the same alternative, but see Remark 1 on this convention. To avoid non-trivialities, we assume that A contains at least two elements.

Formally, a CBCS is a pair (A, P), where A is a non-empty set of alternatives with $|A| \geq 2$ and P is a mapping

$$P : A \times A \to [0, 1]$$

that is *balanced* (Falmagne, 1985, Definition 4.9):

$$P(a, b) = 1 - P(b, a) \tag{1}$$

for all $a, b \in A$. We call such P a *binary choice probability (BCP)* . If $a \neq b$ then $P(a, b)$ is interpreted as the probability (frequency) with which the decision-maker selects a when (given repeated opportunities of) choosing between a or b. The balancedness condition excludes the possibility of abstention — choices are forced by assumption. It also implies that

$$P(a, a) = \frac{1}{2} \tag{2}$$

for any $a \in A$.

Remark 1. Let $\Gamma_A = \{(a, a) \mid a \in A\}$ and let $\mathcal{D}_A = (A \times A) \setminus \Gamma_A$. By defining BCPs on the domain $A \times A$, rather than $\mathcal{D}_A$, we follow an established (though not universal) convention in the literature. In particular, this definition puts our analysis on the same footing as that of Fishburn (1973) and Tversky and Russo (1969) so our results can be directly compared with theirs. However, this convention does raise some interpretive and technical issues. On the interpretive side, there is the question of whether or not probabilities of the form $P(a, a)$ should be accorded behavioural significance. One might refrain from doing so, as most authors who treat binary choice as a special case of multinomial choice model appear, implicitly, to do[2]; or else one could assume that there exist two replicas of each $a \in A$, as in Smith (1976), so that subjects can be offered binary choices between two identical alternatives. In the latter case, (2) becomes a substantive behavioural assumption, provided one can operationalise the idea of choosing the "first" replica over the "second".[3] In the former case, one must be careful that assumption (2) does not introduce unwarranted restrictions on choice behaviour via axioms imposed on P — an issue to which we return in Section 2.4.

If A is finite, we fix an enumeration $A = \{a_1, a_2, \ldots, a_n\}$ so that P may be described by the matrix

$$\boldsymbol{P} = \begin{bmatrix} P(a_1, a_1) & P(a_1, a_2) & \cdots & P(a_1, a_n) \\ P(a_2, a_1) & P(a_2, a_2) & \cdots & P(a_2, a_n) \\ \vdots & \vdots & \ddots & \vdots \\ P(a_n, a_1) & P(a_n, a_2) & \cdots & P(a_n, a_n) \end{bmatrix}. \tag{3}$$

This matrix satisfies $\boldsymbol{P} + \boldsymbol{P}^T = \mathbf{1}$, where $\mathbf{1}$ is a matrix with 1 in every cell. We use $\boldsymbol{P}_{i,\cdot}$ to denote the ith row of $\boldsymbol{P}$ and $\boldsymbol{P}_{\cdot,j}$ to denote the jth column.

Given a CBCS, (A, P), it is natural to impute the following *weak stochastic preference relation* on A: for any $a, b \in A$,

$$a \succsim^P b \quad \Leftrightarrow \quad P(a, b) \geq \frac{1}{2}. \tag{4}$$

(The superscript appended to $\succsim$ emphasises that the binary relation is derived from P.) In other words, a is "weakly stochastically preferred" to b iff the decision-maker chooses a over b at least half of the time. This is equivalent, given (1), to $P(a, b) \geq P(b, a)$. For convenience, we will sometimes refer to $\succsim^P$ simply as the decision-maker's "stochastic preferences".

The asymmetric and symmetric parts of $\succsim^P$ are denoted $\succ^P$ and $\sim^P$ respectively, and satisfy

$$a \succ^P b \quad \Leftrightarrow \quad P(a, b) > \frac{1}{2} \tag{5}$$

and

$$a \sim^P b \quad \Leftrightarrow \quad P(a, b) = \frac{1}{2}.$$

Note that $\succsim^P$ is complete by construction (i.e., for any $a, b \in A$, not necessarily distinct, either $a \succsim^P b$ or $b \succsim^P a$) but $\succsim^P$ need not be transitive. It will be transitive iff (A, P) satisfies *weak stochastic transitivity (WST)* — see Definition 33 in the Appendix.

Fishburn (1973) introduces another useful binary relation that can be defined from P:[4]

$$a \succsim_0^P b \quad \Leftrightarrow \quad P(a, c) \geq P(b, c) \quad \text{for any } c \in A. \tag{6}$$

(Once again, the superscript appended to $\succsim_0$ emphasises its dependence on P.) It is clear that $\succsim_0^P$ is transitive, though it need not be complete. Fishburn (1973, Theorem 1), proves that it is complete iff P satisfies a condition he calls *weak independence* (Definition 40 in the Appendix), which is logically independent of WST (Fishburn, 1973, Figure 1). Moreover, as we will demonstrate (Lemma 12), even if P satisfies weak stochastic transitivity *and* weak independence, so that $\succsim^P$ and $\succsim_0^P$ are both weak orders, it need not be the case that $\succsim^P = \succsim_0^P$. A necessary and sufficient condition for these two binary relations to coincide is the following well-known property:

Definition 2. A CBCS satisfies **weak substitutability** iff the following holds for any $a, b, c \in A$:

$$P(a, b) \geq \frac{1}{2} \quad \Rightarrow \quad P(a, c) \geq P(b, c). \tag{7}$$

Lemma 3. *A CBCS satisfies $\succsim^P = \succsim_0^P$ iff it satisfies weak substitutability.*

---

1 Some authors, such as Suppes, Krantz, Luce, and Tversky (1989), use the term "closed" rather than "complete".

2 See, for example, Fishburn (1998, pp.277 and 284).

3 Assumption (2) is in the spirit of Yellott's (1977) "invariance under uniform expansions of the choice set". (See also Krantz, 1964, p.146.) However, it is not, strictly speaking, implied by Yellott's principle, which applies only to choice sets with at least two distinct alternatives.

4 Properly speaking, Fishburn (1973, p.337), introduces the following *conditional decisiveness* relation: for any $a, b \in A$,

$$a \hat{\succ}_0^P b \quad \Leftrightarrow \quad P(a, c) > P(b, c) \quad \text{for some } c \in A.$$

Defining

$$a \succsim_0^P b \quad \Leftrightarrow \quad (b, a) \notin \hat{\succ}_0^P$$

gives (6). Note that $\succsim_0^P$ is complete iff $\hat{\succ}_0^P$ is asymmetric.