Econ Theory (2013) 53:9–25
DOI 10.1007/s00199-009-0482-9

SYMPOSIUM

Expected utility theory from the frequentist perspective

Tai-Wei Hu

Received: 4 July 2008 / Accepted: 10 June 2009 / Published online: 25 June 2009

Abstract We present an axiomatization of expected utility from the frequentist perspective. It starts with a preference relation on the set of infinite sequences with limit relative frequencies. We consider three axioms parallel to the ones for the von Neumann–Morgenstern (vN–M) expected utility theory. Limit relative frequencies correspond to probability values in lotteries in the vN–M theory. This correspondence is used to show that each of our axioms is equivalent to the corresponding vN–M axiom in the sense that the former is an exact translation of the latter. As a result, a representation theorem is established: The preference relation is represented by an average of utilities with weights given by the relative frequencies.

Keywords Objective probability · Expected utility theory · Frequentist theory of probability · Decision theory

JEL Classification D80 · D81

1 Introduction

We study the von Neumann–Morgenstern (vN–M) expected utility theory from the frequentist perspective of probability. von Neumann and Morgenstern (1944, p. 19)

I am very grateful to Kalyan Chatterjee for his guidance and support. I am also indebted to Jeff Kline for his comments. I received many helpful comments from Mamoru Kaneko that significantly improved the paper.

T.-W. Hu (✉)
Department of Economics, Penn State University, 608 Kern Graduate Building,
University Park, PA 16802-3306, USA
e-mail: tuh129@psu.edu

emphasized this perspective for expected utility theory, but they did not explicitly formulate it in their development of expected utility theory. In the literature after them, the notion of probability in expected utility theory is not considered from the frequentist viewpoint.[1] Expected utility theory based on the frequentist theory will be crucial when we take experiences seriously into account, such as in inductive game theory (Kaneko and Kline 2008). In this paper, we present a frequentist axiomatization of expected utility. This gives a frequentist foundation for expected utility theory.

Although the frequentist interpretation is traced back to the middle of the nineteenth century (Gillies 2000, chap 5), the modern frequentist theory began with von Mises (1981).[2] He attempted to construct a formal system based on infinite sequences of outcomes from repetitive experiments. He gives two requirements for such a sequence:

(i) it has a well-defined limit relative frequency for each outcome;
(ii) it is random in the sense that it has no pattern generated by a finite rule.

A sequence satisfying these two requirements is called a *collective* in von Mises (1981); however, he did not succeed in finding a rigorous definition of the second requirement. Wald (1938) gave the first rigorous definition of randomness, but Ville (1939) showed that it was not yet satisfactory. In the recent literature, a satisfactory definition is emerging (see Downey et al. 2006 for a recent survey).

We reconsider the vN–M expected utility theory from the frequentist perspective of probability, and adopt collectives as the objects of preferences. Requirement (ii) is inessential for our axiomatization in that the main results are not affected regardless of whether or not we impose requirement (ii) in our system. In this paper, we focus on a system without requirement (ii); thus, a collective in this paper is an infinite sequence satisfying requirement (i). We will discuss the results when requirement (ii) is incorporated in Sect. 5.2. Thus, our approach is capable of interpreting probability values in expected utility theory as generated by a well-defined random mechanism or as frequencies regardless of such random mechanisms.

Our approach starts with a preference relation over infinite sequences of outcomes satisfying requirement (i), and we propose three axioms on the preference relation. Our main result shows that our axioms correspond to the vN–M axioms. This correspondence is based on the translation mapping a collective to a lottery having the same probability values as the frequencies in that collective. With this translation, we show that the two axiomatic systems are equivalent. Here, we emphasize that the underlying structures of the two systems are still different, but that the above translation allows us to compare them.

We use the equivalence result to obtain a representation theorem: It states that the preference relation over collectives is represented by the long-run average criterion. This representation theorem, together with the above equivalence theorem, gives the frequentist foundation for expected utility theory.

The rest of the paper is organized as follows. A review of the vN–M expected utility theory is given in Sect. 2. Our axioms, the equivalence and the representation theorems

[1] See Barbera et al. (1998) for a survey on expected utility theory.

[2] In the literature, there are alternatives to the frequentist theory (see Weatherford 1982 or Gillies 2000 for a survey of interpretations of probability); namely, the classical theory, the subjective theory, and the logical theory.

are presented in Sect. 3. A variant system based on finite sequences is presented in Sect. 4. We discuss our results and some possible extensions in Sect. 5. Proofs of the main theorems and lemmas are given in Sect. 6.

2 Expected utility theory

Here, we give a small summary of the vN–M expected utility theory. Consider a finite set of *outcomes* $X = \{x_1, \ldots, x_n\}$. The set of *lotteries* over X is given as

$$\Delta(X) = \left\{ p \in (K[0,1])^n : \sum_{x \in X} p_x = 1 \right\},$$

where $K[0, 1]$ is either the set of real numbers from 0 to 1 ($K = \mathbb{R}$) or the set of rational numbers from 0 to 1 ($K = \mathbb{Q}$). While our main results hold for both cases, the latter case ($K = \mathbb{Q}$) will be used in Sect. 4. An element $p \in \Delta(X)$ is denoted as $p = (p_x)_{x \in X} = (p_{x_1}, p_{x_2}, \ldots, p_{x_n})$.

A *preference relation* $\precsim^P$ is a binary relation over $\Delta(X)$. Given this relation, the *indifference relation* $\sim^P$ and the *strict preference relation* $\prec^P$ are defined as follows: For any $p, q \in \Delta(X)$,

$$p \sim^P q \quad \text{if and only if} \quad p \precsim^P q \text{ and } q \precsim^P p; \tag{1}$$

$$p \prec^P q \quad \text{if and only if} \quad p \precsim^P q \text{ and not } q \precsim^P p. \tag{2}$$

In the vN–M theory, the concept of a *compound lottery* plays an important role; it is the convex combination of the form $ap + (1-a)q$, where $p,\ q \in \Delta(X)$ are two lotteries and $a \in K[0, 1]$ is a number. For a comparison with the corresponding concept in our frequentist approach, we present the operation of taking a convex combination as

$$(p, q, a) \mapsto ap + (1-a)q. \tag{3}$$

We will refer to this mapping when we introduce the corresponding operation in our approach.

The vN–M theory has the following three axioms:

EU1 $\precsim^P$ is a complete and transitive binary relation.

EU2 For all $p,\ q,\ r \in \Delta(X)$, if $p \prec^P q \prec^P r$, then there is some $a \in K[0, 1]$ such that $q \sim^P ap + (1-a)r$.

EU3 For all $p,\ q,\ r \in \Delta(X)$ and $a \in K[0, 1]$ with $a \neq 0$,

$$ap + (1-a)r \precsim^P aq + (1-a)r \quad \text{if and only if} \quad p \precsim^P q.$$

The representation theorem for expected utility is given in Theorem 2.1, where K can be either $\mathbb{R}$ or $\mathbb{Q}$. The proof is a small variant of the standard expected utility theorem, which can be found in Fishburn (1970, pp 112–115).

Theorem 2.1 (Expected utility) *A preference relation* $\precsim^P$ *satisfies EU1–EU3 if and only if there exists a function* $h : X \to K$ *such that for all* $p, q \in \Delta(X)$,

$$p \precsim^P q \Leftrightarrow \sum_{x \in X} p_x h(x) \leq \sum_{x \in X} q_x h(x).$$

3 A frequentist axiomatization

In this section, we give an axiomatization of expected utility from the frequentist perspective. We consider a preference relation over infinite sequences over X satisfying the requirement (i) stated in Sect. 1. We will give three axioms for expected utility in our frequentist approach, and show that those axioms are equivalent to the vN–M axioms under some translation between the two approaches. Using this equivalence result, we obtain a representation of the preference relation by the long-run average criterion.

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$. For any $p = (p_x)_{x \in X} \in \Delta(X)$, a sequence $\xi = (\xi_0, \xi_1, \ldots) \in X^{\mathbb{N}} = \prod_{t \in \mathbb{N}} X$ is called a *p-sequence* iff

$$\lim_{T \to \infty} \frac{|\{t : 0 \leq t \leq T - 1,\ \xi_t = x\}|}{T} = p_x \quad \text{for each } x \in X. \tag{4}$$

That is, the limit relative frequency of x in a p-sequence is p_x for each outcome x. Then, the set of *collectives* over X is defined to be

$$\Omega_X = \{\xi \in X^{\mathbb{N}} : \xi \text{ is a } p\text{-sequence for some } p \in \Delta(X)\}.$$

Collectives here are defined using only requirement (i) mentioned in Sect. 1. The case with requirement (ii) will be discussed in Sect. 5.2.

We consider a preference relation $\precsim$ over the set of collectives Ω_X. The indifference and strict parts, denoted by $\sim$ and $\prec$, respectively, of $\precsim$ are defined in the same manner as in (1) and (2). We will formulate three axioms on the preference relation, using similar ideas to those behind the vN–M axioms.

In our approach, we need an operation corresponding to the compound lottery operation (3). For this purpose, we introduce the *shuffle operator*. Let $\xi = (\xi_0, \xi_1, \ldots)$ and $\zeta = (\zeta_0, \zeta_1, \ldots)$ be two collectives in Ω_X, and let $\nu = (\nu_0, \nu_1, \ldots)$ be an infinite binary sequence in $\{0, 1\}^{\mathbb{N}}$. The collectives ξ, ζ correspond to the lotteries p, q, and ν corresponds to the probability weight a in (3). The shuffle operator is expressed as

$$(\xi, \zeta, \nu) \mapsto \xi \oslash_\nu \zeta. \tag{5}$$

Formally, it is defined by the following equations:

$$(\xi \oslash_\nu \zeta)_0 = (1 - \nu_0)\xi_0 + \nu_0 \zeta_0 \quad \text{for } t = 0, \tag{6}$$

$$(\xi \oslash_\nu \zeta)_t = (1 - \nu_t)\xi_{t - f^\nu(t)} + \nu_t \zeta_{f^\nu(t)} \quad \text{for } t > 0, \tag{7}$$

Table 1 Illustration of the shuffle operator

ν	0	1	0	0	1	1	1	0	...
ξ	ξ_0		ξ_1	ξ_2				ξ_3	...
ζ		ζ_0			ζ_1	ζ_2	ζ_3		...
$\xi \oslash_\nu \zeta$	ξ_0	ζ_0	ξ_1	ξ_2	ζ_1	ζ_2	ζ_3	ξ_3	...

where $f^\nu(t) = \sum_{s=0}^{t-1} \nu_s$ is the number of occurrences of 1's in the initial segment of ν with length t. To illustrate this definition, consider Table 1.

The first line in Table 1 describes the weight sequence ν; in the second line, the elements from ξ appear in the places where ν has value 0; in the third line, the ones from ζ appear in the places where ν has value 1; and those two lines are combined into the shuffled sequence in the bottom. If we put $\nu = (0, 1, 0, 1, 0, 1, 0, 1, \ldots)$, $\xi = (x, y, x, y, x, y, \ldots)$, and $\zeta = (y, x, y, x, y, x, \ldots)$ in Table 1, then the shuffled sequence $\xi \oslash_\nu \zeta$ becomes $(x, y, y, x, x, y, y, x, \ldots)$.

In the expected utility theory as in Sect. 2, it is assumed that a compound lottery is reduced to a lottery in $\Delta(X)$ using convex combination. In our framework, we need to show a corresponding reduction, which is given in the following lemma. Its proof is given in Sect. 6.

Lemma 3.1 *Let $\xi, \zeta \in X^{\mathbb{N}}$, $\nu \in \{0, 1\}^{\mathbb{N}}$, and let $p, q \in \Delta(X)$, $a \in [0, 1]$. Suppose that ξ is a p-sequence, ζ is a q-sequence, and ν is an $(a, 1-a)$-sequence in $\{0, 1\}^{\mathbb{N}}$. Then, $\xi \oslash_\nu \zeta$ is an $(ap + (1-a)q)$-sequence.*

Here, an $(a, 1-a)$-sequence is one with limit relative frequencies a and $1-a$ for 0 and 1, respectively.

Now we are ready to present our axioms on a preference relation $\precsim$ over Ω_X.

- **A1** $\precsim$ is a complete and transitive binary relation.
- **A2** For all ξ, ζ, η in Ω_X, if $\xi \prec \zeta \prec \eta$, then there is a number a in $K[0, 1]$ and an $(a, 1-a)$-sequence ν in $\{0, 1\}^{\mathbb{N}}$ such that $\zeta \sim \xi \oslash_\nu \eta$.
- **A3** For all ξ, ζ, η in Ω_X and all $(a, 1-a)$-sequences ν^1, ν^2 in $\{0, 1\}^{\mathbb{N}}$ with $a \in K[0, 1]$ and $a \neq 0$, $\xi \oslash_{\nu^1} \eta \precsim \zeta \oslash_{\nu^2} \eta$ if and only if $\xi \precsim \zeta$.

Axioms A1–A3 are parallel to EU1–EU3. Perhaps, Axiom A3 needs more comments, which will be given after our theorems.

Now we give a comparison between our axiomatic system and the vN–M system. For this comparison, we use the mapping ψ from Ω_X to $\Delta(X)$:

$$\psi(\xi) = \left(\lim_{T\to\infty} \frac{|\{t : 0 \le t \le T-1,\ \xi_t = x\}|}{T} \right)_{x \in X} \quad \text{for each } \xi \in \Omega_X. \tag{8}$$

The mapping ψ specifies the relative frequencies of outcomes in a collective. We can translate our system $(\Omega_X, \precsim)$ into the vN–M system $(\Delta(X), \precsim^P)$ using the mapping ψ if for all $\xi,\ \zeta \in \Omega_X$,

$$\xi \precsim \zeta \quad \text{if and only if } \ \psi(\xi) \precsim^P \psi(\zeta). \tag{9}$$

This means that the preference relation $\precsim$ over Ω_X can be reduced to the preference relation $\precsim^P$ over $\Delta(X)$, in which case $\precsim$ depends only on relative frequencies.

The following theorem states, first, that our axioms A1 and A3 guarantee the above reduction. Also, it states that under condition (9) each of our axioms corresponds precisely to its parallel axiom in the vN–M theory. Its proof will be given in Sect. 6.

Theorem 3.1 (Frequentist translation)

(a) *Suppose that the preference relation $\precsim$ over Ω_X satisfies A1 and A3. Then there is a preference relation $\precsim^P$ over $\Delta(X)$ satisfying (9).*
(b) *Suppose that $\precsim$ over Ω_X and $\precsim^P$ over $\Delta(X)$ satisfy (9). Then,*
(b1) *$\precsim$ satisfies A1 if and only if $\precsim^P$ satisfies EU1.*
(b2) *$\precsim$ satisfies A2 if and only if $\precsim^P$ satisfies EU2.*
(b3) *$\precsim$ satisfies A3 if and only if $\precsim^P$ satisfies EU3.*

Thus, our axiomatic system is a faithful translation of the vN–M system within our frequentist framework. It justifies the frequentist perspective taken by von Neumann and Morgenstern (1944, p. 19), while their theory is neutral to interpretations of probability.

Theorem 3.1 allows us to obtain our representation theorem. Its proof will be given in Sect. 6.

Theorem 3.2 (Frequentist axiomatization of expected utility) *A preference relation $\precsim$ satisfies A1–A3 if and only if there exists a function $h : X \to K$ (which we call a representing utility function[3] of $\precsim$) such that for all $\xi, \zeta \in \Omega_X$,*

$$\xi \precsim \zeta \Leftrightarrow \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} \leq \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\zeta_t)}{T}. \tag{10}$$

Theorems 3.1 and 3.2 manifest the parallelism between our theory and the vN–M theory. By this parallelism, we can regard our theory as a frequentist version of the vN–M expected utility theory.

An additional comment on the particular form of (10) is still needed; why does the representation take the long-run average criterion? More precisely, where is the hidden assumption for the equal treatment of every element in a collective $\xi = (\xi_0, \xi_1, \xi_2, \ldots)$? Axiom A3 is the driving force of this equal treatment: In Axiom A3, ν^1 and ν^2 are required only to have the same relative frequencies, and the orders of elements in those sequences can be changed. In fact, the main cause is hidden in the present definition (6) and (7) of the shuffle operator. It is crucial in the present proof of Theorem 3.2. In Sect. 5.2, an alternative formulation of the shuffle operator is discussed, but some difficulty is pointed out.

[3] As in the vN–M theory, representing utility functions are determined by the preference relation to be unique up to positive linear transformations.

4 Axiomatization with finite sequences

In this section we show that the framework of Sect. 3 can be restricted to finite sequences, but we can keep the same axiomatization of expected utility with a small modification.

Given a finite set X of outcomes, the set of finite sequences over X is denoted by $X^{<\mathbb{N}}$. Each finite sequence is written as $\sigma = (\sigma_0, \ldots, \sigma_{t-1})$. The length t of σ is denoted by $|\sigma|$.

Consider a preference relation $\precsim^F$ over $X^{<\mathbb{N}}$. We give four axioms, AF1–AF4, on $\precsim^F$; the first three are parallel to Axioms A1–A3, and the last axiom is additional to deal with finite sequences.

The main difference from the previous axiomatization is that the shuffle operator needs to be defined for finite sequences. Another difference is the preference comparison between finite sequences of different lengths, which is taken care of by the new axiom, AF4. Here, the lengths of finite sequences may be different, and this fact causes difficulties with the shuffle operator and preference comparisons. To avoid these difficulties, we replicate input sequences of the shuffle operator so that they have same lengths. For this purpose, we give one definition: For finite sequence σ, $\sigma^t = (\sigma, \sigma, \ldots, \sigma)$ is the sequence obtained from replicating σ t times. This new sequence preserves the relative frequencies of the original sequence σ.

The finite shuffle operator is formally defined as follows: Given two sequences $\sigma, \tau \in X^{<\mathbb{N}}$ and a finite binary sequence $\rho \in \{0,1\}^{<\mathbb{N}}$, they are replicated into $\sigma' = \sigma^{|\tau||\rho|}$, $\tau' = \tau^{|\sigma||\rho|}$, and $\rho' = \rho^{|\tau||\sigma|}$, respectively. Then, those replicated sequences have the same length $|\tau| \cdot |\sigma| \cdot |\rho|$ and preserve the relative frequencies of the original sequences. The *finite-shuffle* of σ and τ with ρ, denoted by $\sigma \oslash_\rho \tau$, has the length $|\tau| \cdot |\sigma| \cdot |\rho|$ and is defined by

$$(\sigma \oslash_\rho \tau)_0 = (1-\rho'_0)\sigma'_0 + \rho'_0\tau'_0 \quad \text{for } t = 0,$$
$$(\sigma \oslash_\rho \tau)_t = (1-\rho'_t)\sigma'_{t-f^{\rho'}(t)} + \rho'_t\tau'_{f^{\rho'}(t)} \quad \text{for } t = 1, \ldots, |\tau| \cdot |\sigma| \cdot |\rho| - 1,$$

where $f^{\rho'}(t) = \sum_{s=0}^{t-1} \rho'_s$.

The finite-shuffle operator is closely related to the shuffle operator defined in Sect. 3: For any finite sequences $\sigma, \tau \in X^{<\mathbb{N}}$, and $\rho \in \{0,1\}^{<\mathbb{N}}$, if $\xi = (\sigma, \sigma, \sigma, \ldots)$, $\zeta = (\tau, \tau, \tau, \ldots)$, and $\nu = (\rho, \rho, \rho, \ldots)$, then $\xi \oslash_\nu \zeta = (\sigma \oslash_\rho \tau, \sigma \oslash_\rho \tau, \sigma \oslash_\rho \tau, \ldots)$. Moreover, Lemma 3.1 can be modified to the present finite framework.

The four axioms on the preference relation $\precsim^F$ are as follows.

AF1: $\precsim^F$ is a complete and transitive binary relation.

AF2: For all $\sigma, \tau, \pi \in X^{<\mathbb{N}}$, if $\sigma \prec^F \tau \prec^F \pi$, then there is a binary sequence $\rho \in \{0,1\}^{<\mathbb{N}}$ such that $\tau \sim^F \sigma \oslash_\rho \pi$.

AF3: For all $\sigma, \tau, \pi \in X^{<\mathbb{N}}$ and $\rho^1, \rho^2 \in \{0,1\}^{<\mathbb{N}}$ that have the same positive relative frequency of outcome 0,

$$\sigma \oslash_{\rho^1} \pi \precsim^F \tau \oslash_{\rho^2} \pi \quad \text{if and only if } \sigma \precsim^F \tau.$$

AF4: For all $\sigma \in X^{<\mathbb{N}}$ and for all $t > 0$, $\sigma \sim^F \sigma^t$.

Axioms AF1–AF3 are very parallel to A1–A3, but AF4 is new. It states that a finite sequence σ is indifferent to any replication of itself. It is the essence of this axiom that the preference relation counts only the relative frequencies but not the lengths. This enables us to compare sequences with different lengths.

The finite sequence version of the mapping ψ in (8) becomes now the mapping $\phi : X^{<\mathbb{N}} \to \Delta(X)$:

$$\phi(\sigma) = \left(\frac{|\{t : 0 \leq t \leq |\sigma| - 1,\ \sigma_t = x\}|}{|\sigma|} \right)_{x \in X}. \tag{11}$$

Translation (9) then becomes for all $\sigma,\ \tau \in X^{<\mathbb{N}}$,

$$\sigma \precsim^F \tau \quad \text{if and only if } \ \phi(\sigma) \precsim^P \phi(\tau), \tag{12}$$

where the preference relation $\precsim^P$ is a binary relation over $\Delta(X)$ with $K = \mathbb{Q}$.

Now, Theorem 3.1 becomes the following.

Theorem 4.1 (a) *Suppose that the preference relation $\precsim^F$ over $X^{<\mathbb{N}}$ satisfies AF1, AF3, and AF4. Then there is a preference relation $\precsim^P$ over $\Delta(X)$ $(K = \mathbb{Q})$ satisfying* (12).

(b) *Suppose that $\precsim^F$ over $X^{<\mathbb{N}}$ and $\precsim^P$ over $\Delta(X)$ satisfy* (12). *Then,*

(b.1) *$\precsim^F$ satisfies AF1 if and only if $\precsim^P$ satisfies EU1.*

(b.2) *$\precsim^F$ satisfies AF2 if and only if $\precsim^P$ satisfies EU2.*

(b.3) *$\precsim^F$ satisfies AF3 if and only if $\precsim^P$ satisfies EU3.*

The proof of Theorem 4.1 is quite parallel to that of Theorem 3.1. The only difference is to use Axiom AF4 to compare finite sequences with different lengths. We do not give the proof in this paper, but it is available upon request.

Then, Theorem 3.2 becomes the following.

Theorem 4.2 *A preference relation $\precsim^F$ satisfies AF1–AF4 if and only if there exists a function $h : X \to \mathbb{Q}$ such that for all $\sigma,\ \tau \in X^{<\mathbb{N}}$,*

$$\sigma \precsim \tau \ \Leftrightarrow \sum_{t=0}^{|\sigma|-1} \frac{h(\sigma_t)}{|\sigma|} \leq \sum_{t=0}^{|\tau|-1} \frac{h(\tau_t)}{|\tau|}.$$

The proof follows exactly the same arguments as that of Theorem 3.2, and is omitted. One difference between the finite approach and the infinite approach is that the randomness requirement (ii) mentioned in Sect. 1 cannot be incorporated in the finite approach. This may be regarded as a demerit from the frequentist perspective of probability. However, from the viewpoint of expected utility theory, this is rather a merit in the sense that it can be applied to finite sequences.

5 Conclusions and remarks

5.1 Conclusions

We gave two axiomatic approaches to expected utility, based on infinite sequences and finite sequences. Theorems 3.1 and 4.1 describe the correspondences between our axiomatic systems and the vN–M system. Theorems 3.2 and 4.2 are frequentist expected utility theorems. As mentioned in Sect. 1, von Neumann and Morgenstern (1944, p. 19), emphasized the frequentist interpretation of probability for their expected utility theory. Our theorems give a justification of their interpretation. Moreover, Theorems 4.1 and 4.2 widened the frequentist perspective to accommodate finite sequences for expected utility theory.

This approach can be used to study game theory, especially for research programs that take into account players' *ex post* experiences. In particular, this approach has the potential to serve a foundation for expected utility in inductive game theory (cf. Kaneko and Kline 2009).

5.2 Remarks

In this subsection, we first give some comments on the technical side of our approach. Then, we discuss a possible extension of our framework and one application.

5.2.1 Randomness requirement

The randomness requirement can be incorporated in our formulation. For this, we define collectives to be infinite sequences satisfying requirements (i) and (ii) in Sect. 1; specifically, we use the definition of random sequences given by Martin-Löf (1966) to formulate requirement (ii). The axioms are formulated in the same way as in Sect. 3, but we restrict the application of the shuffle operator to "independent" random sequences. We can show that Theorem 3.1 and Theorem 3.2 hold in this framework as well.[4]

5.2.2 Shuffle operator

We formulate the shuffle operator to reflect compound lotteries, and our formulation is justified by Lemma 3.1. One alternative operator is the following: Let $\xi,\ \zeta \in X^{\mathbb{N}}$, and let $\nu \in \{0, 1\}^{\mathbb{N}}$; define $\xi \ominus_\nu \zeta$ by setting

$$(\xi \ominus_\nu \zeta)_t = (1 - \nu_t)\xi_t + \nu_t \zeta_t \quad \text{for all } t.$$

However, this simpler operator does not work for Theorems 3.1 and 3.2. The choice of ν may affect the frequency of the shuffled sequence: Lemma 3.1 is no longer valid with this shuffle operator.

[4] The precise definitions and results are available upon request.

If the randomness requirement is incorporated, Lemma 3.1 is recovered, but the present author has not yet succeeded in finding a proof for Theorem 3.1 using this shuffle operator. The main difficulty is that with this alternative operator, in general the input sequences cannot be recovered from the resulting sequence and the weight sequence, while one can always recover them with the shuffle operator defined in Sect. 3.

5.2.3 A Frequentist definition of subjective probability

This paper gives the frequentist version of vN–M expected utility theory. By extending our framework, it is possible to formulate an axiomatic system for a frequentist definition of subjective probability. The subjective probability values thus obtained are interpreted as the personal estimation of the long-run frequencies in a random sequence encountered in the decision maker's problem. We suspect that the axioms in Anscombe and Aumann (1963) can be used to formulate such an axiomatic system.

5.2.4 Application to game theory

Our results are used in Hu (2009), which studies strategic unpredictable behavior from the frequentist perspective. There, a collective game that consists of infinite repetitions of a finite two-person zero-sum game is considered, and a play in the collective game is an infinite sequence of joint actions from the two players. The long-run average criterion is applied to specify the payoffs in that paper, and our axiomatic system provides a foundation for that application.

6 Proofs

We give proofs of the main results in Sect. 3 in this section. In the first subsection, we will provide two more lemmas, together with the proof of Lemma 3.1. In the second subsection, we give the proofs of Theorems 3.1 and 3.2.

6.1 Preliminary lemmas

First we introduce the function $L^{\xi,x} : (\mathbb{N} - \{0\}) \to \mathbb{N}$ that counts the number of occurrences of outcome $x \in X$ in initial segments of a collective $\xi \in \Omega_X$:

$$L^{\xi,x}(T) = |\{t : 0 \leq t \leq T-1, \xi_t = x\}| \quad \text{for all } T > 0.$$

This function will simplify the notations in our proofs. Now we give the proof of Lemma 3.1.

Proof of Lemma 3.1: The key to this proof is to notice the following equation:

$$L^{\xi \oslash_\nu \zeta,x}(T) = L^{\xi,x}(L^{\nu,0}(T)) + L^{\zeta,x}(L^{\nu,1}(T)) \quad \text{for each } T \text{ and for each } x. \quad (13)$$

It states that the number of occurrences of outcome x in an initial segment of the shuffled sequence with length T is the sum of those numbers in the initial segments of the input sequences with lengths specified by the numbers of occurrences of 0 and 1 in the initial segment of the weighting sequence with length T. It is straightforward to check its valid using the definition of the shuffle operator, Eqs. (6) and (7). This equation can be used directly to prove the lemma for $a \in (0, 1)$ as follows: For all T large enough, $L^{\nu,0}(T) > 0$ and $L^{\nu,1}(T) > 0$ (actually, $\lim_{T\to\infty} L^{\nu,0}(T) = \infty = \lim_{T\to\infty} L^{\nu,1}(T)$), and so

$$\begin{aligned}\lim_{T\to\infty} \frac{L^{\xi\oslash_\nu\zeta,x}(T)}{T} &= \lim_{T\to\infty}\left(\frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)}\frac{L^{\nu,0}(T)}{T} + \frac{L^{\zeta,x}(L^{\nu,1}(T))}{L^{\nu,1}(T)}\frac{L^{\nu,1}(T)}{T}\right)\\ &= \lim_{T\to\infty}\frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)}\lim_{T\to\infty}\frac{L^{\nu,0}(T)}{T}\\ &\quad + \lim_{T\to\infty}\frac{L^{\zeta,x}(L^{\nu,1}(T))}{L^{\nu,1}(T)}\lim_{T\to\infty}\frac{L^{\nu,1}(T)}{T} = ap_x + (1-a)q_x.\end{aligned}$$

For the case $a = 1$, the above argument does not work because

$$\lim_{T\to\infty}\frac{L^{\nu,1}(T)}{T} = 1 - a = 0, \tag{14}$$

and so $\lim_{T\to\infty} L^{\nu,1}(T)$ may be a finite number. However, because

$$\lim_{T\to\infty}\frac{L^{\nu,0}(T)}{T} = a = 1,$$

we still have

$$\lim_{T\to\infty}\frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)} = p_x. \tag{15}$$

Now, (14) and $L^{\zeta,x}(L^{\nu,1}(T)) \le L^{\nu,1}(T)$ imply that

$$\lim_{T\to\infty}\frac{L^{\zeta,x}(L^{\nu,1}(T))}{T} = 0.$$

Using (13), we have that

$$\begin{aligned}\lim_{T\to\infty}\frac{L^{\xi\oslash_\nu\zeta,x}(T)}{T} &= \lim_{T\to\infty}\left(\frac{L^{\xi,x}(L^{\nu,0}(T))}{L^{\nu,0}(T)}\frac{L^{\nu,0}(T)}{T}\right)\\ &\quad + \lim_{T\to\infty}\frac{L^{\zeta,x}(L^{\nu,1}(T))}{T} = ap_x = p_x.\end{aligned}$$

The case for $a = 0$ is completely symmetric. □

To prove part (a) of Theorem 3.1, we use the mathematical induction on the number of outcomes that appear in the collective. In the arguments, we decompose a collective into two subsequences according to a subset of outcomes A and use Axiom A3 to obtain the result. Here, we introduce a decomposition operator that will simplify those arguments. Let $A \subset X$ be a subset and let $\xi \in X^{\mathbb{N}}$. Define $\nu^{\xi,A}$ as $\nu_t^{\xi,A} = 1$ if $\xi_t \in A$ and $\nu_t^{\xi,A} = 0$ otherwise, which indicates whether a particular element in ξ belongs to A or not. To decompose ξ according to A via $\nu^{\xi,A}$, we introduce another function θ^ν. For any $\nu \in \{0, 1\}^{\mathbb{N}}$, define the function θ^ν as follows:

$$\theta^\nu(0) \ \text{ is the least } t' \ \text{ such that } \nu_{t'} = 1; \tag{16}$$

$$\theta^\nu(t+1) \ \text{ is the least } t' \ \text{ such that } t' > \theta^\nu(t) \ \text{ and } \ \nu_{t'} = 1. \tag{17}$$

θ^ν records the places where ν has value 1. Notice that $\theta^\nu(t)$ may not be well-defined for all $t \in \mathbb{N}$. It is easy to check that it is well-defined for all t if and only if ν has infinitely many 1's.

Now, we can form the subsequence ξ^A of ξ obtained by eliminating elements in ξ that are not in A, using the function $\theta^{\nu^{\xi,A}}$: The sequence ξ^A is defined as ($x^0 \in X$ is a fixed outcome)

(1) if $\lim_{T\to\infty} |\{t : 0 \le t \le T-1, \ \nu_t^{\xi,A} = 1\}| = \infty$, then let $\xi_t^A = \xi_{\theta^{\nu^{\xi,A}}(t)}$ for all $t \in \mathbb{N}$;
(2) if $\lim_{T\to\infty} |\{t : 0 \le t \le T-1, \ \nu_t^{\xi,A} = 1\}| = K < \infty$, then let $\xi_t^A = \xi_{\theta^{\nu^{\xi,A}}(t)}$ for all $t = 0, \dots, K-1$, and let $\xi_t^A = x^0$ for all $t \ge K$.

If infinitely many elements in ξ are in A, then ξ^A is a subsequence of ξ; otherwise, $\xi_t^A = x^0$ for all t large enough. It is straightforward to check that $\xi^{\overline{A}} \oslash_{\nu^{\xi,A}} \xi^A = \xi$ ($\overline{A}$ denotes the complement of A). That is, ξ can be decomposed into two disjoint subsequences ξ^A and $\xi^{\overline{A}}$ (provided that it has infinite elements belonging to A and $\overline{A}$) such that ξ^A is a sequence over A and $\xi^{\overline{A}}$ is a sequence over $\overline{A}$. Moreover, if ξ is a collective, so are ξ^A and $\nu^{\xi,A}$. Their relative frequencies can be found using conditional probability. The following lemma summarizes these properties.

Lemma 6.1 *Let $p \in \Delta(X)$. Suppose that ξ is a p-sequence. We have*

(1) *$\nu^{\xi,A}$ is an $(1-p_A, p_A)$-sequence, where $p_A = \sum_{x\in A} p_x$.*
(2) *if $p_A > 0$, then ξ^A is a p^A-sequence, where $p_x^A = \frac{p_x}{p_A}$ if $x \in A$ and $p_x^A = 0$ otherwise.*

Proof (1) It is straightforward to check that $L^{\nu^{\xi,A},1}(T) = \sum_{x\in A} L^{\xi,x}(T)$ for all $T > 0$. Thus,

$$\lim_{T\to\infty} \frac{L^{\nu^{\xi,A},1}(T)}{T} = \sum_{x\in A} \lim_{T\to\infty} \frac{L^{\xi,x}(T)}{T} = \sum_{x\in A} p_x = p_A.$$

(2) We first deal with outcomes not in A. For each $x \notin A$, by definition (here part (1) of the definition applies because $p_A > 0$) of ξ^A, $\xi_t^A \neq x$ for all $t \in \mathbb{N}$. Thus,

$$\lim_{T\to\infty} \frac{L^{\xi^A,x}(T)}{T} = 0 = p_x^A.$$

Now we consider outcomes in A. Since $p_A > 0$, $\lim_{T\to\infty} L^{\nu^{\xi,A},1}(T) = \infty$ by part (1) of this lemma. As a result, $\theta^{\nu^{\xi,A}}(t)$ is well-defined for all $t \in \mathbb{N}$. Because $\theta^{\nu^{\xi,A}}$ is strictly increasing by Eq. (17), we have that $\lim_{T\to\infty} \theta^{\nu^{\xi,A}}(T) \to \infty$.

We claim that

$$\lim_{T\to\infty} \frac{T}{\theta^{\nu^{\xi,A}}(T-1)+1} = p_A. \tag{18}$$

By construction, $\theta^{\nu^{\xi,A}}(T-1)$ is the place in $\nu^{\xi,A}$ where the Tth occurrence of 1 takes place. Thus, there are exactly T occurrences of 1's in the first $\theta^{\nu^{\xi,A}}(T-1)+1$ elements in $\nu^{\xi,A}$, and so

$$L^{\nu^{\xi,A},1}(\theta^{\nu^{\xi,A}}(T-1)+1) = T \quad \text{for all } T \geq 1. \tag{19}$$

Hence, by (19), the sequence $\{\frac{T}{\theta^{\nu^{\xi,A}}(T-1)+1}\}_{T=1}^{\infty}$ is a subsequence of $\{\frac{L^{\nu^{\xi,A},1}(T)}{T}\}_{T=1}^{\infty}$. By part (1), the latter sequence has limit p_A, and so the former sequence has the same limit. This validates (18).

For any $x \in A$,

$$\begin{aligned}
\lim_{T\to\infty} \frac{L^{\xi^A,x}(T)}{T} &= \lim_{T\to\infty} \frac{L^{\xi,x}(\theta^{\nu^{\xi,A}}(T-1)+1)}{\theta^{\nu^{\xi,A}}(T-1)+1} \frac{\theta^{\nu^{\xi,A}}(T-1)+1}{T} \\
&= \lim_{T\to\infty} \frac{L^{\xi,x}(\theta^{\nu^{\xi,A}}(T-1)+1)}{\theta^{\nu^{\xi,A}}(T-1)+1} \lim_{T\to\infty} \frac{\theta^{\nu^{\xi,A}}(T-1)+1}{T} \\
&= p_x \frac{1}{\lim_{T\to\infty} \frac{T}{\theta^{\nu^{\xi,A}}(T-1)+1}} = \frac{p_x}{p_A} = p_x^A.
\end{aligned} \tag{20}$$

It is straightforward to check that $L^{\xi^A,x}(T) = L^{\xi,x}(\theta^{\nu^{\xi,A}}(T-1)+1)$ and this gives the first equality in (20). The last equality in (20) comes from (18) and the fact that the limit of the inverses of a sequence is equal to the inverse of its limit. □

We end this subsection with a lemma that links expected values to long-run averages. This lemma is the key step to obtain the long-run average criterion (10).

Lemma 6.2 *Let $h : X \to \mathbb{R}$ be any function. Suppose that $\xi \in X^{\mathbb{N}}$ is a p-sequence for some $p \in \Delta(X)$. Then*

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \sum_{x\in X} p_x h(x).$$

Proof For any $x' \in X$, $h(x') = \sum_{x \in X} c_{x'}(x)h(x)$, where $c_{x'}(x) = 1$ if $x = x'$ and $c_{x'}(x) = 0$ otherwise. We first show that for all $x^0 \in X$,

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{c_{x^0}(\xi_t)}{T} = \sum_{x\in X} p_x c_{x^0}(x) = p_{x^0}. \tag{21}$$

For each $T \in \mathbb{N}$,

$$\sum_{t=0}^{T-1} c_{x^0}(\xi_t) = |\{t : 0 \le t \le T-1,\ \xi_t = x^0\}|.$$

Since ξ is a p-sequence, it follows from (4) that

$$\lim_{T\to\infty} \frac{|\{t : 0 \le t \le T-1,\ \xi_t = x^0\}|}{T} = p_{x^0}.$$

Thus, (21) holds. Therefore, we have

$$\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{\sum_{x\in X} c_x(\xi_t)h(x)}{T}$$
$$= \sum_{x\in X} \lim_{T\to\infty} h(x) \sum_{t=0}^{T-1} \frac{c_x(\xi_t)}{T} = \sum_{x\in X} h(x)p_x = \sum_{x\in X} p_x h(x).$$

□

6.2 Proofs of main theorems

Proof of Theorem 3.1: **(a)** We claim that if $\precsim$ satisfies A1 and A3, then for any $p \in \Delta(X)$ and any p-sequences ξ and ζ, $\xi \sim \zeta$. Given this claim, it is straightforward to construct a preference relation $\precsim^P$ that satisfies (9).

First we give some notations. For any $x \in X$, we use $\mathbf{x}$ to denote the sequence such that for all $t \in \mathbb{N}$, $\mathbf{x}_t = x$. Also, for any $p \in \Delta(X)$, we define $S(p) = \{x \in X : p_x > 0\}$, that is, the set of outcomes in the support of p.

We prove the claim by induction on the number of outcomes in $S(p)$.

The basis Suppose that $|S(p)| = 1$. Let $S(p) = \{x\}$. The difficulty here arises when a p-sequence still contains other outcomes than x.

Let ξ be a p-sequence. Consider $\xi^{\overline{\{x\}}}$ and $\nu^{\xi,\overline{\{x\}}}$. By construction, we have that $\xi = \mathbf{x} \oslash_{\nu^{\xi,\overline{\{x\}}}} \xi^{\overline{\{x\}}}$, and, by Lemma 6.1, $\nu^{\xi,\overline{\{x\}}}$ is an $(1,0)$-sequence. By A3,

$$\xi = \mathbf{x} \oslash_{\nu^{\xi,\overline{\{x\}}}} \xi^{\overline{\{x\}}} \precsim \mathbf{x} \oslash_{\mathbf{0}} \xi^{\overline{\{x\}}} = \mathbf{x} \quad \text{if and only if } \ \mathbf{x} \precsim \mathbf{x},$$

and since $\mathbf{x} \sim \mathbf{x}$ by A1, $\xi \sim \mathbf{x}$ (here, $\mathbf{0} = (0, 0, 0, \ldots.)$ is the sequence consisting of 0's). Thus, for any an p-sequence ξ, $\xi \sim \mathbf{x}$, and so by A1, for any two p-sequences ξ and ζ, we have $\xi \sim \zeta$.

The inductive step Our induction hypothesis is that for any $p \in \Delta(X)$ with $|S(p)| \leq k$, $k \geq 1$, any two p-sequences are indifferent.

Consider any $q \in \Delta(X)$ with $|S(q)| = k + 1$ and any q-sequences ξ and ζ. We will show that $\xi \sim \zeta$.

Let $y \in S(q)$, and let $A = S(q) - \{y\}$. First we show that there are q-sequences ξ' and ζ' that consists of only outcomes in $S(q)$ such that $\xi \sim \xi'$ and $\zeta \sim \zeta'$. Moreover, we show that $\xi' \sim \zeta'$, and this will give us the desired result.

Notice that $\xi = \xi^{\overline{A}} \oslash_{\nu^{\xi,A}} \xi^A$ ($q_{\overline{A}} > 0$ and $q_A > 0$), and, by Lemma 6.1, $\nu^{\xi,A}$ is an $(1 - q_A, q_A)$-sequence. Because $\xi^{\overline{A}}$ is a $p^{\{y\}}$-sequence ($p_x^{\{y\}} = 1$ if $x = y$ and $p_x^{\{y\}} = 0$ otherwise), by the induction hypothesis, $\xi^{\overline{A}} \sim \mathbf{y}$.

Define $\xi' = \mathbf{y} \oslash_{\nu^{\xi,A}} \xi^A$. Then, by A3 and $\xi^{\overline{A}} \sim \mathbf{y}$, we have $\xi' \sim \xi$. Similarly, we define ζ' as $\mathbf{y} \oslash_{\nu^{\zeta,A}} \zeta^A$, and, using the same arguments, we have $\zeta' \sim \zeta$.

By Lemma 6.1, both ξ^A and ζ^A are q^A-sequences. By the induction hypothesis, $\xi^A \sim \zeta^A$. Notice that ξ' can be written as $\xi^A \oslash_{\nu^{\xi,\overline{A}}} \mathbf{y}$ and ζ' can be written as $\zeta^A \oslash_{\nu^{\zeta,\overline{A}}} \mathbf{y}$. Since both $\nu^{\xi,\overline{A}}$ and $\nu^{\zeta,\overline{A}}$ are $(q_A, 1 - q_A)$-sequences, it follows from A3 that $\xi' \sim \zeta'$

Because $\xi' \sim \xi$, $\zeta' \sim \zeta$, and $\xi' \sim \zeta'$, it follows from A1 that $\xi \sim \zeta$. This completes our inductive step. By mathematical induction, we have proved our claim.

Now, we define $\precsim^P$ to be such that $p \precsim^P q$ if and only if $\xi \precsim \zeta$ for some p-sequence ξ and some q-sequence ζ. The above claim shows that $\precsim^P$ is well-defined and satisfies condition (9).

(b) Suppose that there is a preference relation $\precsim^P$ over $\Delta(X)$ that satisfies (9).

(b.1) ($\Rightarrow$) Suppose that $\precsim$ satisfies A1, i.e, it is complete and transitive. First we show that $\precsim^P$ is transitive. Let $p, q, r \in \Delta(X)$ be such that $p \precsim^P q$ and $q \precsim^P r$. There are $\xi, \zeta, \eta \in \Omega_X$ such that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence. Then, by (9), $\xi \precsim \zeta$ and $\zeta \precsim \eta$. Since $\precsim$ is transitive, it follows that $\xi \precsim \zeta$. Hence, by (9), $p \precsim^P r$. The proof for completeness is similar.

($\Leftarrow$) Suppose that $\precsim^P$ satisfies EU1. Let $\xi, \zeta, \eta \in \Omega_X$ be such that $\xi \precsim \zeta$ and $\zeta \precsim \eta$. Let $p, q, r \in \Delta(X)$ be such that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence. It then follows, from (9), that $p \precsim^P q$ and $q \precsim^P r$. Since $\precsim^P$ is transitive, it follows that $p \precsim^P r$. Hence, by (9), $\xi \precsim \eta$. The proof for completeness is similar.

(b.2) ($\Rightarrow$) Suppose that $\precsim$ satisfies A2. We show that $\precsim^P$ satisfies EU2 by considering lotteries $p, q, r \in \Delta(X)$ that satisfy $p \prec^P q \prec^P r$. Let ξ be a p-sequence, ζ be a q-sequence, and η be a r-sequence. Then, by (9), $\xi \prec \zeta \prec \eta$. By A2, there exists a number $a \in K[0, 1]$ and an $(a, 1 - a)$-sequence $\nu \in \{0, 1\}^{\mathbb{N}}$ such that $\zeta \sim \xi \oslash_\nu \eta$. By Lemma 3.1, $\xi \oslash_\nu \eta$ is an $(ap + (1 - a)r)$-sequence. Thus, by (9) and $\zeta \sim \xi \oslash_\nu \eta$, $q \sim^P ap + (1 - a)r$.

($\Leftarrow$) Suppose that $\precsim^P$ satisfies EU2. Suppose that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence, and suppose that $\xi \prec \zeta \prec \eta$. Then, by (9), $p \prec^P q \prec^P r$. By EU2, there is a number $a \in K[0, 1]$ such that

$q \sim^P ap + (1-a)r$. Pick any $(a, 1-a)$-sequence ν, then $\xi \oslash_\nu \eta$ is an $(ap + (1-a)r)$-sequence. Hence, by (9) and $q \sim^P ap + (1-a)r$, $\eta \sim \xi \oslash_\nu \eta$.

(b.3) $(\Rightarrow)$ Suppose that $\precsim$ satisfies A3. Recall that EU3 has two directions.

First we show that 'if' directions. Let $p, q, r \in \Delta(X)$, and let $a \in K[0, 1]$ with $a \neq 0$. Suppose that $p \precsim^P q$. Let ξ be a p-sequence ξ and let ζ be a q-sequence. Then $\xi \precsim \zeta$. Pick any r-sequence η and pick any $(a, 1-a)$-sequence $\nu \in \{0, 1\}^{\mathbb{N}}$. By Lemma 3.1, $\xi \oslash_\nu \eta$ is an $(ap + (1-a)r)$-sequence and $\zeta \oslash_\nu \eta$ is an $(aq + (1-a)r)$-sequence. By A3, $\xi \precsim \zeta$ implies that

$$\xi \oslash_\nu \eta \precsim \zeta \oslash_\nu \eta.$$

Thus, $ap + (1-a)r \precsim^P aq + (1-a)r$.

For the 'only if' direction, suppose that $ap + (1-a)r \precsim^P aq + (1-a)r$. Let ξ be a p-sequence, let ζ be a q-sequence, let η be a r-sequence, and let ν be an $(a, 1-a)$-sequence. By Lemma 3.1, $\xi \oslash_\nu \eta$ is an $(ap + (1-a)r)$-sequence and $\zeta \oslash_\nu \eta$ is an $(aq + (1-a)r)$-sequence. Hence, by (9), $\xi \oslash_\nu \eta \precsim \zeta \oslash_\nu \eta$. By A3, $\xi \precsim \zeta$, and so $p \precsim^P q$.

$(\Leftarrow)$ Suppose that $\precsim^P$ satisfies EU3. Again, A3 has two directions.

For the 'if' direction, suppose ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence. Also, suppose that $\xi \precsim \zeta$. Then, $p \precsim^P q$. For any $a \in K[0, 1]$ with $a \neq 0$, using EU3, we have that $ap + (1-a)r \precsim^P aq + (1-a)r$. By Lemma 3.1, for any $(a, 1-a)$-sequences ν^1 and ν^2, $\xi \oslash_{\nu^1} \eta$ is an $(ap + (1-a)r)$-sequence and $\zeta \oslash_{\nu^2} \eta$ is an $(aq + (1-a)r)$-sequence. Therefore, we have $\xi \oslash_{\nu^1} \eta \precsim \zeta \oslash_{\nu^2} \eta$.

Consider the 'only if' direction. Suppose that ξ is a p-sequence, ζ is a q-sequence, and η is a r-sequence, and suppose that ν^1 and ν^2 are $(a, 1-a)$-sequences with $a \in K[0, 1]$ and $a \neq 0$. If $\xi \oslash_{\nu^1} \eta \precsim \zeta \oslash_{\nu^2} \eta$, it follows from Lemma 3.1 and (9) that $ap + (1-a)r \precsim^P aq + (1-a)r$. Then, by EU3, $p \precsim r$. Hence, we have $\xi \precsim \zeta$. □

Proof of Theorem 3.2: There are two directions in this theorem. In part (a) of the proof, we show that 'if' direction, and in part (b) of the proof, we show the other direction.

(a) We show that existence of a representing utility function h of $\precsim$ implies that $\precsim$ satisfies A1–A3. Let $p, q \in \Delta(X)$. By Lemma 6.2, if ξ is a p-sequence, then $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \sum_{x\in X} p_x h(x)$. Similarly, if ζ is a q-sequence, $\lim_{T\to\infty} \sum_{t=0}^{T-1} \frac{h(\zeta_t)}{T} = \sum_{x\in X} q_x h(x)$. Therefore, by (10),

$$\xi \precsim \zeta \quad \text{if and only if } \sum_{x\in X} p_x h(x) \leq \sum_{x\in X} q_x h(x). \tag{22}$$

Now, define $\precsim^P$ over $\Delta(X)$ as $p \precsim^P q$ if and only if $\sum_{x\in X} p_x h(x) \leq \sum_{x\in X} q_x h(x)$. It is straightforward, using (22), to check that (9) holds for $\precsim$ and $\precsim^P$. By Theorem 2.1, it follows that $\precsim^P$ satisfies EU1–EU3. Then, by Theorem 3.1, we have that $\precsim$ satisfies A1–A3.

(b) Conversely, suppose that $\precsim$ satisfies A1–A3. By Theorem 3.1, there exists a preference relation $\precsim^P$ over $\Delta(X)$ that satisfies EU1–EU3 and satisfies condition (9). By Theorem 2.1, axioms EU1–EU3 hold if and only if there is a function $h : X \to K$ such that for all $p,\ q \in \Delta(X)$,

$$p \precsim^P q \quad \text{if and only if} \quad \sum_{x \in X} p_x h(x) \leq \sum_{x \in X} q_x h(x).$$

By Lemma 6.2, if ξ is a p-sequence, then

$$\lim_{T \to \infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} = \sum_{x \in X} p_x h(x). \tag{23}$$

Therefore, for any p-sequence ξ and any q-sequence ζ,

$$\xi \precsim \zeta \quad \text{if and only if} \quad p \precsim^P q \quad \text{if and only if} \quad \sum_{x \in X} p_x h(x) \leq \sum_{x \in X} q_x h(x),$$

which, by (23), is equivalent to

$$\lim_{T \to \infty} \sum_{t=0}^{T-1} \frac{h(\xi_t)}{T} \leq \lim_{T \to \infty} \sum_{t=0}^{T-1} \frac{h(\zeta_t)}{T}.$$

Thus, h is a representing utility function of $\precsim$. □

References

Anscombe, F.J., Aumann, R.: A definition of subjective probability. Ann Math Stat **34**, 199–205 (1963)

Barbera, S., Hammond, P.J., Seidl, C.: Handbook of Utility Theory: Principles, vol. 1. Dordrecht: Kluwer (1998)

Downey, R., Hirschfeldt, D., Nies, A., Terwijn, S.: Calibrating randomness. Bull Symb Logic **12**, 411–491 (2006)

Fishburn, P.C.: Utility Theory for Decision Making. New York: Wiley (1970)

Gillies, D.: Philosophical Theories of Probability. London: Routlegde (2000)

Hu, T.-W.: Complexity and Mixed Strategy Equilibria. http://taiweihu.weebly.com/research.html (2009)

Kaneko, M., Kline, J.J.: Inductive game theory: a basic scenario. J Math Econ **44**, 1332–1363 (2008)

Kaneko, M., Kline, J.J.: Transpersonal understanding through social roles, and emergence of cooperation. Tsukuba University, Department of Social Systems and Management Discussion Paper Series, No. 1228 (2009)

Martin-Löf, P.: The definition of random sequences. Inform Control **9**, 602–619 (1966)

Ville, J.: Ětude Critique de la Concept du Collectif. Paris: Gauthier-Villars (1939)

von Mises, R.: Probability, Statistics, and Truth, 2nd revised English edition, New York, Dover (1981)

von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton: Princeton University (1944)

Wald, A.: Die widerspruchsfreiheit des kollektivsbegriffes. In: Wald: Selected Papers in Statistics and Probability, pp. 25–41, 1955. New York: McGraw-Hill (1938)

Weatherford, R.: Philosophical Foundations of Probability Theory. London: Routledge and Kegan Paul (1982)