

Discussion Papers No. 221, June 1998
Statistics Norway, Research Department

John K. Dagsvik

Choice among Lotteries when Preferences are Stochastic

Abstract:

This paper discusses the problem of specifying probabilistic models for choices (strategies) with uncertain outcomes. The point of departure is an extension of the axiom system of the von Neumann-Morgenstern Expected utility theory to the case when the preferences are stochastic. This extended axiom system is combined with Luce Choice Axiom; "Independence from Irrelevant Alternatives", and imply a particular choice model that contains the Luce model as a special case. An additional invariance assumption is subsequently proposed that yields a complete characterization of the mathematical structure of the model.

Keywords: Random tastes, choice among lotteries, random utility models, bounded rationality, probabilistic choice models, independence from irrelevant alternatives.

JEL classification: C25, D11, D81.

Acknowledgement: Criticism and comments by Rolf Aaberge and Steinar Strøm are highly appreciated.

Address: John K. Dagsvik, Statistics Norway, Research Department. E-mail: john.dagsvik@ssb.no

Discussion Papers

comprises research papers intended for international journals or books. As a preprint a Discussion Paper can be longer and more elaborated than a usual article by including intermediate calculation and background material etc.

Abstracts with downloadable postscript files of
Discussion Papers are available on the Internet: <http://www.ssb.no>

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
P.O. Box 1260
N-2201 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

1. Introduction

This paper develops a class of probabilistic choice models for choice experiments in which the outcomes are uncertain to the agent. This means that the agent's response to the same choice situation (with uncertain outcomes) is assumed to be governed by a probability mechanism, and so in general he exhibits inconsistencies. By now there is a huge literature on stochastic choice models with certain outcomes. (For a summary of these models, see Suppes et al. (1989), ch. 17, and Anderson et al. (1992), ch. 2.) In fact, it was empirical observations of inconsistencies, dating back to Thurstone (1927a,b), that lead to the study of probabilistic theories in the first place. Thurstone argued that one reason for the observed inconsistencies is that the agent has difficulties with assessing the precise value (to him) of the choice objects. While probabilistic models for certain outcomes have been studied and applied extensively in psychology and economics it seems that there has been little interest in developing corresponding models for choice with uncertain outcomes (cf. Machina, 1985). This is rather curious since one would expect that if an agent has problems with rank ordering alternatives with certain outcomes he would certainly find it difficult to choose among gambles. The importance of developing theoretically justified stochastic choice models in this context has been accentuated in two recent papers, Harless and Camerer (1994) p. 1287, and Hey and Orme (1994). For example, Hey and Orme, p.p. 1321-1322, summarize their view as follows:

"Our results suggest quite strongly that the truth is *not* going to be found along this deterministic choice route, unless some account is taken of the errors. There is clearly a problem of identifying the underlying "true" model because of these errors—indeed it could be argued that the lack of significance for some of the top-level functionals (deterministic non-expected utility functionals) for some of the subjects in our study could simply result from this noise, ...".

In the next paragraph they conclude:

"....., we are tempted to conclude by saying that our study indicates that behavior can be reasonably well modelled (to what might be termed a 'reasonable approximation') as 'Expected utility plus noise'. Perhaps we should now spend some time thinking about the noise, rather than about even more alternatives to expected utility?"

The point of departure in this paper is to combine a version of the von Neumann-Morgenstern Expected Utility Theory with some of the ideas that have emerged in the literature on discrete choice models with certain outcomes to obtain a theoretical rationale for a probabilistic choice model for uncertain outcomes. Specifically, we introduce a version of a von Neumann-Morgenstern axiom system subject to stochastic preferences, suitably defined. The notion of stochastic preferences is

defined as follows: Suppose the agent faces n replications of a binary choice experiment in which lotteries r and s (say) are presented in each experiment. If the fraction of times lottery r is chosen over s is larger than 0.5 when n is "large", lottery r is said to be preferred over s . From the axioms it follows immediately that one can represent preferences by the expected utility of the respective lotteries. However, this result is incomplete in our context since it does not provide a link between the expected utilities and the choice probability for each replication of the experiment. These probabilities are essential for establishing the link between theory and the corresponding empirical model. To this end, we propose a version of Luce's Choice Axiom, cf. Luce (1959) also known as "Independence from Irrelevant Alternatives" assumption (IIA). As a special case of the model that follows from these axioms one obtains the *Strict expected utility model* proposed by Becker et al. (1963a) and Luce and Suppes (1965). However, these authors provide no theoretical justification for their model other than the property that it contains Luce model (for certain outcomes) as a special case. Becker et al. (1963a) and Luce and Suppes (1965) also consider other types of stochastic choice models for uncertain outcomes. Consequently, the present approach provides a rationalization of the model proposed by Becker et al. (1963a), and Luce and Suppes (1965). However, in this paper we push the theory a step further by postulating a particular invariance principle which enables us to derive important functional form properties.

The paper is organized as follows: Section 2 contains a discussion of the choice axioms and the derivation of the implied structural choice probabilities. In Section 3 a random utility representation is discussed. Here, we demonstrate that the choice model is compatible with a random utility representation.¹ In the final section we demonstrate that in particular choice settings, a special case of the choice model developed under assumptions made in Section 2 has the same formal structure as a model for choice under perfect certainty but with choice sets that are latent to the analyst.

Although the model developed in this paper is a stochastic version of the expected utility model, it is easily realized how this model in some cases could be extended to corresponding stochastic non-expected utility models. This is the case for the Rank Dependent Expected Utility Model, (cf. Allais (1979), Quiggin (1982), Yaari (1987), Chew, Karni and Safra (1987), and the Subjective Expected Utility Model, such as Edwards (1962), and Kahneman and Tversky (1979). What all of these non-expected utility models have in common is that the conditional probabilities for the respective outcomes given the respective choices are replaced by a function of these probabilities.

¹ In Dagsvik (1995) a somewhat related approach was proposed to derive probabilistic choice models for uncertain outcomes. This approach was, however, not based on behavioral axioms to the same extent as the present paper.

2. An axiom system for probabilistic choice among lotteries

We mentioned above that in microeconomic theory the tradition is to assume that the consumer has a utility function that allows him to rank the alternatives in a consistent and unambiguous manner when faced with identical choice settings. This approach has been criticized by psychologists and others, see for example Thurstone (1927a,b), Luce (1959), Tversky (1969), who argue that in choice settings people often experience uncertainty and inconsistency. That is, they have difficulties with assessing the precise (subjective) value of the alternatives and consequently the choice outcomes in identical choice settings may vary across settings. To account for this empirical evidence the psychologists have developed probabilistic choice models. In the psychological choice literature one has traditionally distinguished between two types of choice models: In the *constant utility* model the decision rule is viewed as stochastic while utility is deterministic (see Luce and Suppes (1965)). Luce model (Luce (1959)) is the most famous example of a constant utility model. Luce derives this model from his choice axiom (IIA) and demonstrates that it implies the existence of a unique (except for a multiplicative constant) scale (constant—or deterministic utility) and the choice probabilities can be expressed as a function of the scale values by a simple formulae. In the *random utility* model, utility is viewed as stochastic (Thurstone) while the decision rule is deterministic. In light of recent work by economists it seems that the difference between these models is only superficial. Specifically, Holman and Marley (cited in Luce and Suppes (1965)) and McFadden (1981) have demonstrated that the most familiar constant utility models such as the Luce model and Tversky's "elimination by aspects" model, Tversky (1972), can both be represented by random utility formulations.

Becker et al. (1963a,b) and Luce and Suppes (1965) extend the Luce model for perfectly certain outcomes to the case with uncertain outcomes by simply specifying the scale values as a function of the corresponding expected utilities. As mentioned above, this approach is somewhat ad hoc from a theoretical point of view and we shall next discuss a possible theoretical foundation for the models proposed by Becker et al. (1963a,b) and Luce and Suppes (1965). First we need some additional notation and terminology.

Let S denote the total set of lotteries and let X denote the set of outcomes which is finite and contains m outcomes.

Definition 1

A preference relation is a binary relation, \succeq , on S that is (i) complete, i.e. for all $r, s \in S$ either $r \succeq s$ or $s \succeq r$, and (ii) transitive, i.e. for all r, s, t , in S , $r \succeq s$ and $s \succeq t$ implies $r \succeq t$.

Definition 2

A real-valued function $V(s)$ on S represents \succeq if for all $r, s \in S$, $r \succeq s$ if and only if $V(r) \geq V(s)$.

In our context we shall, as mentioned above, allow the agent to have random preferences in the sense that, if the agent faces several replications of a specific choice experiment he may choose different lotteries each time. The reasons for this is that he may have difficulty with evaluating the proper value (to him) of the respective lotteries.

Consider now the following choice setting: The agent faces n replications of a binary choice experiment in which lotteries r and s (say) are presented in each experiment. Since the agent has stochastic preferences he may choose lottery r in some replications and lottery s in the remaining ones. Let $\hat{P}_n(r, s)$ be the fraction of the n replications where lottery r is chosen over s . Evidently, it seems natural to say that r is preferred over s if $\hat{P}_n(r, s) > \hat{P}_n(s, r)$ when n is "large". Since

$$P(r, s) = \text{plim}_{n \rightarrow \infty} \hat{P}_n(r, s),$$

where $P(r, s)$ is the theoretical probability, it follows that $P(r, s) > P(s, r)$ if and only if $P(r, s) > 0.5$. Accordingly, the argument above provides a motivation for the following definition:

Definition 3

For $r, s \in S$, lottery r is strictly preferred to s if and only if $P(r, s) > 0.5$. If $P(r, s) = 0.5$, then r is indifferent to s .

Thus Definition 3 introduces a binary relation, \succeq , where $r \succ s$ means that r is strictly preferred to s , while $r \sim s$ means that r is indifferent to s . However, the relation is not necessary a preference relation.

In the following we shall assume, as is customary, that the agent's information about the chances of the different realizations can be represented by probabilities;

$$g_s \equiv (g_s(1), g_s(2), \dots, g_s(m))$$

where $g_s(k)$ is the probability of outcome k , $k \in X$, if lottery s is chosen. Let G denote the set of simple probability measures on the algebra of all subsets of the set of outcomes, $X \equiv \{1, 2, \dots, m\}$. Let $g_1, g_2 \in G$. The mixed lottery, $\alpha g_1 + (1 - \alpha)g_2$, $\alpha \in [0, 1]$, is a lottery in G yielding the consequence $\alpha g_1(k) + (1 - \alpha)g_2(k)$, $k \in X$. Here we assume that the lotteries $\alpha \beta g_1 + (1 - \alpha \beta)g_2$ and $\beta[\alpha g_1 + (1 - \alpha)g_2] + (1 - \beta)g_2$, $\alpha, \beta \in [0, 1]$, are indifferent. This property is known as the axiom of reduction of compound lotteries, cf. Luce and Raiffa (1957).

Assumption A1 (Continuity)

For all $g_1, g_2, g_3 \in G$, if $g_1 \succ g_2$ and $g_2 \succ g_3$, then there exists $\alpha, \beta \in (0, 1)$ such that

$$\alpha g_1 + (1 - \alpha)g_3 \succ g_2$$

and

$$g_2 \succ \beta g_1 + (1 - \beta)g_3.$$

Assumption A2 (Independence)

For all $g_1, g_2, g_3 \in G$, and all $\alpha \in [0, 1]$, if $g_1 \succeq g_2$, then

$$\alpha g_1 + (1 - \alpha)g_3 \succeq \alpha g_2 + (1 - \alpha)g_3.$$

The Assumptions A1 and A2 provide fundamental underpinnings of the von Neumann-Morgenstern theory for decision under uncertainty. The next theorem is a slightly modified version of Theorem 2.4 in Karlin and Schmeidler (1991).

Theorem 1

Let \succeq be the binary relation given in Definition 3. The following two conditions are equivalent:

- (i) \succeq is a preference relation satisfying Assumptions A1 and A2.

(ii) There exists a function $W: G \rightarrow R$, such that for any $g \in G$;

$$W(g) = \varphi \left(\sum_{k \in X} u(k) g(k) \right)$$

where $u: X \rightarrow R$, is a function that is unique up to a positive affine transformation and $\varphi: R \rightarrow R$ is a strictly increasing function.

Theorem 1 is a slightly modified version of the von Neumann-Morgenstern Expected Utility Theorem. In the conventional deterministic theory for decision under uncertainty the result that W , given in Theorem 1, (ii) represents \succeq on G yields a complete result insofar as it enables us to express the agents preferences by an index W as a function of the agents "information" represented by the probabilities associated with the lotteries. In the present context, however, the binary relation \succeq is defined in terms of choice probabilities.² We can only conclude so far that $P(r, s) \geq 0.5$ if and only if $W(g_r) \geq W(g_s)$. In the special case with certain outcomes the statement; $P(r, s) \geq 0.5$ if and only if $W(g_r) \geq W(g_s)$, is equivalent to Definition 16, p. 333, in Luce and Suppes (1965).

We are so far, however, unable to specify a choice model for *each* replication. Consequently, it remains to develop additional theory so as to be able to ascertain precisely how the choice probabilities $\{P(r, s)\}$ for each replication depends on $(W(g_r), W(g_s))$. This is crucial for establishing a relationship between the theoretical concepts introduced above and a structural probabilistic model that is applicable for empirical modelling and analysis. More generally, in order to establish a complete theory we need to develop a structural specification for the choice probabilities in the general multinomial case, i.e., we need to specify how these probabilities depend on $\{W(g_r), r \in S\}$. This is the issue to be discussed next. To this end we need some additional notation.

The agent is now assumed to face a set of lotteries B , where B is a subset of S . Thus B is the set of lotteries that are feasible to the agent (possibly agent specific). Let $P(s; B)$ denote the probability that the agent shall choose lottery s when B is the choice set of lotteries, $s \in B \subset S$. For sets such that $A \subset B \subset S$, let

$$P(A; B) \equiv \sum_{s \in A} P(s; B).$$

² Note that it is implicit in the treatment in this paper that the agent does not know the distribution of the random variables that affect his preferences. Consequently, he is unable to take account of this distribution when forming expectations.

The interpretation is that $P(A;B)$ is the probability that the agent shall choose a lottery within A when B is the choice set.

Assumption A3 (IIA)

For any $A \subset B \subset S$, $P(A; B) \in (0, 1)$, and

$$P(s; B) = P(s; A) P(A; B).$$

Assumption A3 was first proposed by Luce (1959). It represents a probabilistic version of rationality in the following sense: Suppose the agent face a set B of feasible lotteries. One may view the agent's choice as if it takes place in two stages. In stage one he selects A from B , where A is a subset of B which contains the most attractive lotteries. In a second stage he selects the preferred lottery from the subset A . It is important to stress that Assumption A3 implies that in the second stage, the agent *only* takes into account the lotteries within A . In other words, the lotteries within $B \setminus A$ are *irrelevant* in the second stage. Thus, the rationality is related to the property that the agent only takes into consideration the lotteries within the presented choice set. As the probability of the first stage choice is $P(A;B)$ and that of the second stage is $P(s;A)$, $P(A;B)P(s;A)$ is the final probability of choosing s . Since Assumption A3 is a probabilistic statement it means that it represents probabilistic rationality in the sense that lotteries outside the second stage choice set A may matter in single choice experiments but will not affect average behavior.

Theorem 2

Assumption A3 holds if and only if

$$P(s; B) = \frac{a(s)}{\sum_{r \in B} a(r)}$$

where $\{a(s), s \in S\}$ are positive scalars that are unique apart from a multiplicative positive constant.

For a proof of Theorem 1, see for example Ben-Akiva and Lerman (1985). Theorem 1 was originally proved by Luce (1959).

Theorem 3

Assume that Assumptions A1, A2 and A3 hold. Then

$$(1) \quad P(s; B) = \frac{h(V(g_s))}{\sum_{r \in B} h(V(g_r))}$$

where

$$(2) \quad V(g_s) = \sum_{k \in X} u(k) g_s(k)$$

and $h: \mathbb{R} \rightarrow \mathbb{R}_+$ is a strictly increasing function. For a given g_s , $h(V(g_s))$ is unique apart from a multiplicative positive constant, and $u: X \rightarrow \mathbb{R}$ is a function that is unique up to a positive affine transformation.

Proof:

Let $B = \{r, s\}$. Then

$$P(s; \{r, s\}) = \frac{a(s)}{a(s) + a(r)} = \frac{1}{1 + a(r)/a(s)}.$$

Thus

$$P(s; \{r, s\}) \geq 0.5 \Leftrightarrow a(s) \geq a(r)$$

and $\{a(s), s \in S\}$ therefore represents \succeq on S . But then, by Theorem 1, $a(s)$ must be a strictly increasing function \tilde{h} (say) of $W(g_s)$. Hence, by Theorem 1

$$a(s) = \tilde{h}(W(g_s)) = \tilde{h}(\varphi(V(g_s))) = h(V(g_s))$$

where $h(y) = \tilde{h}(\varphi(y))$, for $y \in \mathbb{R}$. Since $a(s)$ is positive h must be positive.

Q.E.D.

The model stated in Theorem 3 was originally proposed by Becker et al. (1963a). However, their only argument to support the model structure is that it contains the Luce model as a special case. They also derive the following nonparametric result.

Corollary 1

Suppose B consists of $n > 1$ lotteries where

$$g_n(j) = \frac{1}{n-1} \sum_{r=1}^{n-1} g_r(j).$$

Then

$$P(n; B) \leq \frac{1}{n}$$

if h is convex. If h is concave, then

$$P(n; B) \geq \frac{1}{n}.$$

If $h(y) = y$ and $u(k) \geq 0$ for $k \in X$, then

$$P(n; B) = \frac{1}{n}.$$

In the case where $h(y) = y$, the model above reduces to the strict expected utility model for uncertain outcomes proposed by Luce and Suppes (1965).

By Theorem 3 we have brought the theory an important step further in that we have established the relationship between the multinomial choice probabilities and the expected utility functional apart from an increasing mapping. It is important to emphasize that we have achieved this result relying entirely on theoretical principles. No ad hoc functional form assumptions have been invoked. We shall next introduce an additional assumption which will enable us to derive restrictions on the mapping h .

To this end consider the following setting. Recall that the original set of outcomes X consisted of m outcomes. Now suppose that X is expanded in the following manner. Each outcome, apart from the first one, is duplicated $w - 1$ times, i.e., to X we add $w - 1$ outcomes with utilities $u(2)$ and outcome probabilities $g_s(2)$, for each $s = 1, 2, \dots, m$. Subsequently, $w - 1$ outcomes are added to X with utilities $u(3)$ and outcome probabilities, $g_s(3)$, and so on. Obviously, the expected utility of lottery s that corresponds to the expanded set X^* (say) equals

$$(3) \quad V^*(g_s) = \sum_{k \in X^*} u(k) g_s(k) = u(1) + \sum_{k \in X^* \setminus \{1\}} (u(k) - u(1)) g_s(k) = u(1) + w \sum_{k \in X \setminus \{1\}} (u(k) - u(1)) g_s(k).$$

Since by Theorem 1, u is only unique up to a positive affine transformation we may without loss of generality choose an enumeration such that $u(1) = \min_{k \in X} u(k) = 0$, so that $V^*(g_s)$ reduces to

$$(4) \quad V^*(g_s) = w \sum_{k \in X \setminus \{1\}} u(k) g_s(k) = w \sum_{k \in X} u(k) g_s(k).$$

Definition 4

Let $w > 1$ be a natural number. By uniform w -expansion of the set of outcomes X we mean the following: Without loss of generality, let $u(1) = \min_{k \in X} u(k) = 0$. Then expand X by adding $w - 1$ outcomes that are copies of outcome k , for $k = 2, 3, \dots, m$.

Let $P_w(s; B)$ denote the probability of choosing lottery s after uniform w -expansion of X .

Assumption A4

Let $\{g_s, s \in S\}$ and $\{\tilde{g}_s, s \in S\}$ be outcome probabilities such that if the corresponding uniformly expanded choice probabilities $P_w(s; B)$ and $\tilde{P}_w(s; B)$ satisfy

$$P_w(s; B) < \tilde{P}_w(s; B)$$

for some $s \in B$, then for any natural number k

$$P_k(s; B) < \tilde{P}_k(s; B).$$

The intuition behind Assumption A4 is as follows: Outcome one serves as a "reference" outcome and it has the property that the expected contribution to the agents' expected utility associated with this outcome is zero, as seen directly from (4). Since the original outcomes are duplicated uniformly the expected utility of lottery s relative to any other lottery will remain unaltered. As stated in Assumption A4, it seems plausible that if the fraction of agents that prefer lottery g_s from B is less than the fraction of agents that prefer lottery \tilde{g}_s from B , then this should be true also after a rescaling of the utility index $\{u(k)\}$.

Theorem 4

Assume that Assumptions A1 to A4 hold. Then, if h is continuous the choice probabilities have the structure

$$P(s; B) = \frac{h(V(g_s))}{\sum_{r \in B} h(V(g_r))}$$

for $s \in B \subset S$, where

$$h(y) = \exp\left(b \frac{(y^\theta - 1)}{\theta}\right),$$

for $y \geq 0$, and $b > 0$, and θ are constants.³

Proof:

Let $B = \{1, 2\}$. Then

$$(5) \quad P_w(1; \{1, 2\}) = F[h(w V_1)/h(w V_2)]$$

where $V_s = V(g_s)$, $s = 1, 2$, and

$$(6) \quad F(y) = \frac{y}{1+y}.$$

By assumption

$$(7) \quad F(h(w V_1)/h(w V_2)) < F(h(w \tilde{V}_1)/h(w \tilde{V}_2))$$

if and only if

$$(8) \quad F(h(k V_1)/h(k V_2)) < F(h(k \tilde{V}_1)/h(k \tilde{V}_2))$$

where $\tilde{V}_s = V(\tilde{g}_s)$, $s = 1, 2$, and k and w are natural numbers. Let $Y_s = k V_s$ and $\tilde{Y}_s = k \tilde{V}_s$. Then (7) and (8) imply that

³ As usual $\frac{y^\theta - 1}{\theta}$ is defined as $\log y$ when $\theta = 0$.

$$(9) \quad F(h(\lambda Y_1)/h(\lambda Y_2)) < F(h(\lambda \tilde{Y}_1)/h(\lambda \tilde{Y}_2))$$

if and only if

$$(10) \quad F(h(Y_1)/h(Y_2)) < F(h(\tilde{Y}_1)/h(\tilde{Y}_2))$$

where λ is a positive rational number. Since any real number can be approximated arbitrarily closely by rational numbers and h is continuous, (9) and (10) must also hold when λ is a positive real number that belongs to some interval. But then the hypothesis of Theorem 14.19 in Falmagne (1985), p. 338, is fulfilled and it follows from Falmagne's discussion on a particular application of his Theorem 14.19, pp. 338-339 that there exists a strictly increasing and continuous function $H(\cdot)$ on \mathbb{R} such that

$$(11) \quad F(h(V_1)/h(V_2)) = H\left(\frac{V_1^\theta - 1}{\theta} - \frac{V_2^\theta - 1}{\theta}\right)$$

where θ is a constant. Let $\tilde{h}(x)$ be given by

$$\log h(V) = \tilde{h}\left(\frac{V^\theta - 1}{\theta}\right)$$

which implies that (11) can be written as

$$(12) \quad \tilde{h}\left(\frac{V_1^\theta - 1}{\theta}\right) - \tilde{h}\left(\frac{V_2^\theta - 1}{\theta}\right) = f\left(\frac{V_1^\theta - 1}{\theta} - \frac{V_2^\theta - 1}{\theta}\right)$$

where $f(x) = \log(F^{-1} H(x))$. But (12) is equivalent to

$$(13) \quad \tilde{h}(x) = \tilde{h}(y) + f(x - y)$$

where x and y belong to some interval I in \mathbb{R} . Evidently, it is possible to find a g such that $V(g) = 1$, which means that I contains zero. With $y = 0$ in (13) we obtain that

$$f(x) = \tilde{h}(x) - \tilde{h}(0).$$

Without loss of generality we can normalize \tilde{h} such that $\tilde{h}(0) = 0$. Hence (13) can be written

$$(14) \quad \tilde{h}(y + z) = \tilde{h}(y) + \tilde{h}(z)$$

for $y, z \in I$. Eq. (14) is the well known variant of a Cauchy equation and has the solution

$$\tilde{h}(y) = b y$$

where b is a constant, see for example Falmagne (1985), p. 82. This completes the proof.

Q.E.D.

Remark

When $\beta = 1$ and $\theta = 0$, the model reduces to the *Strict Expected Utility* model proposed by Luce and Suppes (1965), p. 360.

3. A random utility representation

As mentioned above, the hypothesis of a random utility index as a representation of preferences dates back to Thurstone (1927). The interpretation of Thurstone's random utility theory is that while the decision rule is deterministic and follows from maximizing utility at each moment, the agent's tastes may fluctuate from one moment to the next in a way that is unpredictable to him. Alternatively, the agent is viewed as being unable to fix a definite (subjective) value of the alternatives.

We shall now answer the question of whether there exists a utility representation which implies choice probabilities as in Theorem 2. In settings where the agent knows the choice sets, Holman and Marley (see Luce and Suppes, 1965, p. 338), McFadden (1973), Yellott (1977) and Strauss (1979) have analyzed the problem of necessary and sufficient conditions for random utility models to satisfy IIA.

The choice probabilities that follow from a random utility model are defined formally by

$$(15) \quad P(A; B) = P\left(\max_{s \in A} U_s = \max_{s \in B} U_s\right)$$

for $A \subset B \subset S$, where $\{U_s, s \in S\}$ are random variables. Let N be the number of lotteries in S . When the joint c.d.f. of (U_1, U_2, \dots, U_N) is specified (15) can, at least in principle, be calculated.

Theorem 5

Assume a random utility model with utilities $\{U_s\}$ given by

$$(16) \quad U_s = \eta_s + \log h(V(g_s))$$

where $\eta_s, s \in S$, are i.i.d. random variables with

$$(17) \quad P(\eta_s \leq y) = \exp(-e^{-y}),$$

$y \in \mathbb{R}$. Then the choice probabilities are given by Theorem 2.

Proof:

When (16) and (17) hold the structure (1) follows readily by straight forward calculus. The proof is also found in standard textbooks of discrete choice theory.

Q.E.D.

Similarly to the case with perfectly certain outcomes, the random utility setting allows us to relax Assumption A3 in an intuitive way.

Assumption A5

The choice probabilities can be expressed as

$$P(s; B) = P\left(v_s + \eta_s = \max_{j \in B} (v_j + \eta_j)\right)$$

where $\{v_s, s \in S\}$ are scalars and $(\eta_1, \eta_2, \dots, \eta_N)$ is a vector of random variables with identical marginal distributions and with joint distribution function that is independent of $\{v_s, s \in S\}$. The corresponding probability measure is positive for all sets with positive Lebesgue measure.

Assumption A5 only states that the decision rule is based upon maximization of random variables that have an additive structure, and it is therefore rather weak. Necessary and sufficient conditions for choice models being generated from random utilities are given by Falmagne (1978).

Theorem 6

Suppose Assumptions A1, A2 and A5 hold. Then

$$v_s = \psi \left(\sum_{k \in X} u(k) g_s(k) \right)$$

where $\psi: \mathbb{R} \rightarrow \mathbb{R}$ is a strictly increasing function.

Proof:

Let $B = \{r, s\}$. Then

$$P(s; \{r, s\}) = P(v_s + \eta_s > v_r + \eta_r) = P(v_s - v_r > \eta_r - \eta_s) = M(v_s - v_r)$$

where $M(\cdot)$ is the c.d.f. of $\eta_r - \eta_s$ that is, by assumption, independent of r and s and also independent of v_r and v_s and is strictly increasing. Since η_r and η_s have the same distribution, $M(\cdot)$ must be symmetric. Hence

$$P(s; \{r, s\}) \geq 0.5 \Leftrightarrow v_s \geq v_r$$

and $\{v_s, s \in S\}$ therefore represents \succeq on S . But then, by Theorem 1, v_s must be a strictly increasing function ψ (say) of $V(g_s)$.

Q.E.D.

From Theorems 5 and 6 we realize immediately that when $\eta_s, s \in S$, are i.i.d. with c.d.f. (17) then we obtain the choice probabilities of Theorem 2.

Theorem 7

Assume Assumptions A1, A2, A4 and A5. Then

$$\psi(y) = \frac{b(y^\theta - 1)}{\theta}$$

for $y \geq 0$, where $b > 0$, and θ are constants.

Proof:

Let

$$M(y) = P(\eta_2 - \eta_1 < \log y)$$

for $y > 0$. With the same notation as in the proof of Theorem 4 it follows that

$$P_w(1; \{1, 2\}) = P\{\eta_2 - \eta_1 < \psi(w V(g_1)) - \psi(w V(g_2))\} = M(\tilde{\psi}(w V(g_1)) / \tilde{\psi}(w V(g_2)))$$

where

$$\tilde{\Psi}(x) = \exp \psi(x).$$

The rest of the proof is completely analogous to the proof of Theorem 4.

Q.E.D.

4. Uncertainty versus aggregation of latent alternatives

Consider now an alternative choice setting. The agent is in this case assumed to be perfectly certain about the outcomes. The universal set S is assumed to have a tree structure. Let B denote the set of feasible lower level "branches" while C_s denote the set of feasible upper level conditional on lower level branch s . Let \tilde{U}_{sj} denote the utility of (s,j) where $j \in C_s, s \in B \subset S$. To the analyst only the choices within B are observable, while the choices within C_s , conditional on the choice of lower level branch s is unobservable. Let κ_s denote the number of alternatives in C_s , and let

$$q_s = \frac{\kappa_s}{\sum_{r \in S} \kappa_r}.$$

Assume moreover that

$$\tilde{U}_{sj} = \log \gamma_s + \varepsilon_{sj}$$

where $\{\gamma_s\}$ are positive deterministic terms, while $\varepsilon_{sj}, j \in C_s, s \in S$ are i.i.d. random variables with joint c.d.f.

$$(18) \quad P\left(\bigcap_{s \in S} \bigcap_{j \in C_s} (\varepsilon_{sj} \leq y_{sj})\right) = \exp\left(-\sum_{s \in S} \left(\sum_{j \in C_s} e^{-y_{sj}/\rho}\right)^\rho\right).$$

The parameter ρ is constrained to $0 < \rho \leq 1$, and has the interpretation

$$\text{corr}(\varepsilon_{sj}, \varepsilon_{si}) = 1 - \rho^2$$

when $j \neq i$, and ε_{sj} and ε_{ri} are independent when $s \neq r$ for all i and j . The c.d.f. (18) is a special case of a Generalized Extreme Value distribution, cf. McFadden (1981), p.p. 228-229. From (18) it follows readily that the choice probabilities $\{\tilde{P}(s; B)\}$, defined by

$$\tilde{P}(s; B) \equiv P\left(\max_{j \in C_s} \tilde{U}_{sj} = \max_{r \in B} \max_{j \in C_r} \tilde{U}_{rj}\right),$$

are given by

$$(19) \quad \tilde{P}(s; B) = \frac{\gamma_s q_s^p}{\sum_{r \in B} \gamma_r q_r^p}$$

for $s \in B$.

Let us now compare this model with a particular extension of the model for uncertain outcomes where the agent also has preferences over lotteries, i.e.,

$$U_s = \log h\left(\sum_{k \in X} u_s(k) g_s(k)\right) + \eta_s$$

where the deterministic utility components $\{u_s(k), k \in X\}$ now depend on the lottery s , $s \in S$.

Assume also that $m = 2$. Without loss of generality we can let $u_s(1) = 0$. If $\eta_s, s \in S$, are i.i.d. random terms with c.d.f. given by (17) and

$$h(y) = \frac{b(y^\theta - 1)}{\theta}$$

with $\theta > 0$, we get from Theorem 5 that

$$(20) \quad P(s; B) \equiv P\left(U_s = \max_{r \in B} U_r\right) = \frac{u_s(2)^\beta g_s(2)^\beta}{\sum_{r \in B} u_r(2)^\beta g_r(2)^\beta}$$

for $s \in B$.

It is intriguing that the mathematical structure of (20) is completely analogous to the structure of (19) when $\{g_s(2)\}$ is replaced by $\{q_s\}$ and $u_s(2)^\beta$ by γ_s . Note also that from the viewpoint of the analyst, q_s may be interpreted as the probability that C_s is non-empty, i.e., that an alternative of "type" s is feasible. By comparing (19) and (20) we conclude that in the choice setting discussed above, and under the hypothesis of rational expectations, the choice probabilities do not depend on whether the agent is uncertain about the outcomes on one hand, or, on the other hand, is perfectly certain about the outcomes, while the analyst is "uncertain" about the "degree of feasibility" of the alternatives within C_s , $s \in B$. Here the degree of feasibility is represented by the fractions $\{q_s\}$.

5. Conclusion

In this paper we have developed a theory of probabilistic choice with uncertain outcomes. By combining the IIA assumption with a particular version of the von Neumann-Morgenstern axiom system we have demonstrated that models proposed earlier in the literature can be given a theoretical rationale. By means of a particular invariance property, we have also demonstrated that one can obtain a characterization of the functional form of the model.

References

- Allais, M. (1979): The so-called Allais Paradox and Rational Decisions Under Uncertainty. In M. Allais and O. Hagen (eds.), *Expected utility hypotheses and the Allais Paradox*, pp. 437-681. D. Reidel Publishing Co., Dordrecht, Holland.
- Anderson, S.P., A. de Palma and J.F. Thisse (1992): *Discrete choice theory of product differentiation*. MIT Press, London.
- Becker, G.M., M.H. De Groot and J. Marschak (1963a): Stochastic Models and Choice Behavior. *Behavioral Sciences*, **8**, 41-55.
- Becker, G.M., M.H. De Groot and J. Marschak (1963b): An Experimental Study of Some Stochastic Models for Wagers. *Behavioral Sciences*, **8**, 199-202.
- Ben-Akiva, M. and S.R. Lerman (1985): *Discrete choice analysis*. MIT Press, Cambridge.
- Chew, C.S., E. Karni and Z. Safra (1987): Empirical Tests of Weighted Utility Theory. *Journal of Mathematical Psychology*, **30**, 55-72.
- Dagsvik, J.K. (1995): Probabilistic Choice Models for Uncertain Outcomes. *Discussion Papers*, no. 141, Statistics Norway.
- Edwards, W. (1962): Subjective Probabilities Inferred from Decisions. *Psychological Review*, **69**, 109-135.
- Falmagne, J.C. (1978): A Representation Theorem for Finite Random Scale Systems. *Journal of Mathematical Psychology*, **18**, 52-72.
- Falmagne, J.C. (1985): *Elements of psychophysical theory*. Oxford University Press, London.
- Harless, D.W. and C.F. Camerer (1994): The Predictive Utility of Generalized Expected Utility Theories. *Econometrica*, **62**, 1251-1290.
- Hey, J.D. and C. Orme (1994): Investigating Generalizations of Expected Utility Theory using Experimental Data. *Econometrica*, **62**, 1291-1326.
- Kahnemann, D. and A. Tversky (1979): Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, **47**, 263-291.
- Karni, E. and D. Schmeidler (1991): Utility Theory with Uncertainty. In W. Hildenbrand and H. Sonnenschein (eds.), *Handbook of Mathematical Economics*, Elsevier Science Publishers B.V., New York.
- Luce, R.D. (1959): *Individual choice behavior: A theoretical analysis*. Wiley, New York.
- Luce, R.D. and H. Raiffa (1957): *Games and decisions*. Wiley, New York.
- Luce, R.D. and P. Suppes (1965): Preference, Utility, and Subjective Probability. In R.D. Luce, R.R. Bush, and E. Galanter (eds.), *Handbook of mathematical psychology, III*, 249-410. Wiley, New York.

Machina, M. (1985): Stochastic Choice Functions Generated from Deterministic Preferences over Lotteries. *Economic Journal*, **92**, 805-824.

McFadden, D. (1973): Conditional Logit Analyses of Qualitative Choice Behavior. In P. Zarembka (eds.), *Frontiers in Econometrics*. Academic Press, New York.

McFadden, D. (1981): Econometric Models of Probabilistic Choice. In C.F. Manski and D. McFadden (eds.), *Structural analysis of discrete data with econometric applications*. MIT Press, London.

Quiggin, J. (1982): A Theory of Anticipated Utility. *Journal of Economic Behavior and Organization*, **3**, 323-343.

Suppes, P., D.H. Krantz, R.D. Luce and A. Tversky (1989): *Foundation of measurement*, Vol. 2. Academic Press, New York.

Strauss, D. (1979): Some Results on Random Utility Models. *Journal of Mathematical Psychology*, **20**, 35-52.

Thurstone, L.L. (1927a): A Law of Comparative Judgment. *Psychological Review*, **34**, 272-286.

Thurstone, L.L. (1927b): Psychological Analysis. *American Journal of Psychology*, **38**, 368-389.

Tversky, A. (1969): Intransitivity of Preferences. *Psychological Review*, **76**, 31-48.

Tversky, A. (1972): Choice-by-Elimination. *Journal of Mathematical Psychology*, **9**, 341-367.

Yaari, M.E. (1987): The Dual Theory of Choice Under Risk. *Econometrica*, **55**, 95-115.

Yellott, J.L. Jr. (1977): The Relationship between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment and the Double Exponential Distribution. *Journal of Mathematical Psychology*, **15**, 109-146.