Relative Probabilities

Frank Heinemann

Goethe-Universität Frankfurt am Main

September, 23rd, 1997

Abstract *In this paper a theory of relative probability measures (RPMs) is developed and related to the standard theory of probability measures. An RPM assigns a non–negative real or infinity to any pair of events. This number should be interpreted as relative probability. An RPM distinguishes possible from impossible events, allows to describe Bayesian updating on null sets, and induces an intuitive notion of independence.*

JEL classification: C 60

Keywords: probability theory; Bayesian updating

Address of the author: Frank Heinemann
Goethe–Universität Frankfurt
Professur für Wirtschaftstheorie
Postfach 81
60 054 Frankfurt a.M.
Germany

FAX: +49 - 69 - 798 250 65
E-mail: fheinema@wiwi.uni–frankfurt.de

1 Introduction

One of the aims of probability theory is to give a formal description of individual beliefs and expectations that can be used to explain the processing of information, resulting changes of beliefs and individually satisfying decisions. Savage (1972 [1954]) has given an axiomatic foundation of subjective probability that led him to model beliefs as probability measures in the definition of Kolmogorov (1933). Information is usually described as knowledge about the occurrence of events. Changes of beliefs are then modeled as a process of Bayesian updating so that posterior beliefs are conditional probabilities.

Kolmogorov's theory of probability has been criticized for its treatment of null events. While those play a minor role in natural sciences, game theoretic solution concepts pay special attention to events that occur with probability zero if the agents follow the proposed strategies. There is an ongoing debate on how to model the rationality of agents[1], and it is an open question, how a rational agent should decide when decision knots are reached that had an a–priori–probability of zero under the presumption that all players are rational. The representation of beliefs by probability measures does not suffice to deal with these issues.

Rényi (1976 [1955, 1956]) developed the notion of a conditional probability space that does not suffer from these shortcomings. Similar structures have been invented by game theorists: Myerson (1986, 1991), Mc Lennan (1989), and Battigalli and Veronese (1996) explored conditional probability systems. Blume, Brandenburger, and Dekel (1991) and Stahl (1995) used lexicographical probabilities. Kohlberg and Reny (1992), Swinkels (1994), and Heinemann (1995a,b) applied relative probabilities.

All of these notions have in common that they allow to derive unique con-

[1] See e.g. Binmore (1992) for an overview, Aumann (1995), or Reny (1995).

ditional probabilities even when the condition is a null set. The basic trick is to define a function that assigns a real to any *pair* of events, while certain axioms guarantee that these values can be interpreted as conditional or relative probabilities.

In this paper I propose to describe individual beliefs as relative probability measures. They assign a non–negative real or infinity to any pair of events. This number stands for the relative probability of the two events. Relative probabilities allow to distinguish possible from impossible events and to describe Bayesian updating on null sets. They give way for a more intuitive notion of independence and they are a powerful tool to describe various axioms for the procession of information.

In section 2 we start with probability measures and show in which respects they are not satisfying. Section 3 introduces relative probability measures. An axiomatic definition is given, some properties and rules of calculus are analyzed. In section 4 we compare relative to absolute probability measures. Section 5 deals with conditional probabilities and section 6 with independence. A conclusion is stated in section 7. An appendix contains all proofs.

2 Probablity Measures

Let W be a nonempty abstract space and $\mathcal{W}$ a σ–field in W. A subset $A \in \mathcal{W}$ is called event, a singular subset $\{z\}$, with $z \in W$ is called elementary event. The empty set $\emptyset$ is an impossible event, W is a sure event.

Definition 1 *A* **probability measure** *on $\mathcal{W}$ is a function μ obeying to the following axioms:*

1. $\mu : \mathcal{W} \to [0, 1]$.

2. $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$.

3. If $\{A_n\}$ *is a sequence of events with* $A_n \searrow \emptyset$ *then* $\mu(A_n) \searrow 0$.

4. $\mu(W) = 1$.

The set of all probability measures on $\mathcal{W}$ is denoted by $\mathcal{M}$. A probability measure μ on $\mathcal{W}$ assigns a number $\mu(A) \in [0,1]$ to any event $A \in \mathcal{W}$. This number stands for the probability of event A. Axiom 2 describes additivity of probabilities, Axiom 3 is the monotone continuity and axiom 4 requires that the sure event has probability 1.

First, note that axiom 3 implies that the impossible event has probability zero, but an event with probability zero does not need to be impossible. An event with probability 1 is said to be almost sure, but it need not be a sure event. A probability measure does neither distinguish between possible and impossible nor between sure and almost sure events.

Definition 2 *Be* $\mu \in \mathcal{M}$ *and* $B \in \mathcal{W}$. *A* **conditional probability measure** *for* μ *under condition* B *is a probability measure* μ_B, *with*

$$\mu_B(A)\,\mu(B) = \mu(A \cap B) \qquad \forall A \in \mathcal{W}.$$

If $\mu(B) > 0$ there is a unique conditional probability measure for μ under B. In this case the conditional probability of event A for μ and B is

$$\mu_B(A) := \mu(A \cap B)/\mu(B).$$

However, if $\mu(B) = 0$ and if the σ–algebra in B induced by $\mathcal{W}$ contains more than two events, there is an infinite number of conditional probability measures for μ under B. Here, *any* probability measure $\tilde{\mu}$ with $\tilde{\mu}(B) = 1$ is a conditional probability measure for μ under B. Independence or continuity assumptions or other considerations can be used to determine conditional probabilities. But, this sort of information is not contained in the probability measure itself. It must be stated in addition to probabilities.

Definition 3 *Be* $\mu \in \mathcal{M}$*. Two* **events** *A and B* **are** *called* **independent of each other***, if*

$$\mu(A)\,\mu(B) = \mu(A \cap B).$$

If $\mu(A) = 0$, then any event is independent of A, even an event $B \supseteq A$ that is implied by A. The latter contradicts our intuition about independent events.

3 Relative Probability Measures

A relative probability measure π assigns a number $\pi(A, B)$ to any ordered pair of events $\langle A, B \rangle$. This number is an expression of the relative frequency of these two events. In order to interpret the function π in this way, it must obey to some axioms which are closely related to the properties of probability measures.

Definition 4 *A* **relative probability measure (RPM)** *on $\mathcal{W}$ is a function π obeying to the following axioms:*

1. $\pi : \mathcal{W} \times \mathcal{W} \to \mathbb{R}_+ \cup \{\infty\}$.

2. *If* $\pi(A \cup B, C) \in \mathbb{R}_+$ *then*

$$\pi(A, C) + \pi(B, C) = \pi(A \cup B, C) + \pi(A \cap B, C).$$

3. If $\pi(A,B) \in \mathbb{R}_+$ *and* $\pi(B,C) \in \mathbb{R}_+$ *then*

$$\pi(A,B)\,\pi(B,C) = \pi(A,C).$$

4. $\pi(A,B) = 0 \quad \Leftrightarrow \quad \pi(B,A) = \infty.$

5. If $\{A_n\}$ *is a sequence of events with* $A_n \searrow \emptyset$ *then* $\pi(A_n \cap B, B) \searrow \pi(\emptyset, B)$.

6. $\pi(\emptyset, A) \in \{0,1\}$.

7. $\pi(\emptyset, W) = 0$.

The set of all relative probability measures on $\mathcal{W}$ is denoted by $\mathcal{P}$.

Interpretation $\pi(A,B) = k$ is interpreted as "event A is k–times as likely as event B." If $\pi(\emptyset, A) = 0$ [1] we say that event A is [im]possible.

Axiom 2 guarantees additivity of relative probabilities. It defines a first rule of calculus. Axioms 3 and 4 define a rule for multiplication. Axiom 4 defines ∞ as the multiplicative inverse of zero. Axiom 5 requires monotone continuity. Axiom 6 demands that each event is either possible or impossible. There are no degrees of possibility beyond probability. Axiom 7 is a non–triviality condition and requires that the sure event is possible.

Note that axioms 3 and 4 imply transitivity of the binary relation "at least as likely as" that is implicitly defined by π.

Given axioms 1 – 6, axiom 7 excludes exactly one function $\pi_\emptyset$, defined by

$$\pi_\emptyset(A,B) = 1 \quad \text{for all } A, B \in \mathcal{W}.$$

We will refer to $\pi_\emptyset$ below. Now, let us state some basic properties, helpful in calculations with relative probabilities:

Lemma 1 *For each $\pi \in \mathcal{P}$ and all $A, B, C, D \in \mathcal{W}$ it is true that*

1. $\pi(A, A) = 1$,
2. *if* $\pi(A, B) \in \mathbb{R}_{++}$ *then* $\pi(A, B) = 1/\pi(B, A)$,
3. *if* $A_n \nearrow B$ *then* $\pi(A_n, B) \nearrow 1$,
4. $\pi(B, W) > 0$ *implies* $\pi(\emptyset, B) = 0$,
5. *if* $\pi(A, B) = \infty$ *and* $\pi(B, C) \neq 0$ *or if* $\pi(A, B) \neq 0$ *and* $\pi(B, C) = \infty$ *then* $\pi(A, B)\, \pi(B, C) = \infty$.

Axioms 1 – 7 suffice to interpret the values of π as relative probabilities. You can calculate with them as intuition suggests. The following section will show the relation of relative to absolute probability measures.

4 Absolute Probabilities

From any RPM π a probability measure $\mu[\pi]$ can be derived that assigns absolute probabilities to all events (see proposition 1). Vice versa, each probability measure can be completed to an RPM (see propositions 2.A and 5).

Proposition 1 *For each $\pi \in \mathcal{P}$ the function $\mu[\pi] : \mathcal{W} \to \mathbb{R} \cup \{\infty\}$, defined by*

$$\mu[\pi](A) := \pi(A, W),$$

is a probability measure on $\mathcal{W}$.

Probability measure $\mu[\pi]$ assigns probabilities relative to the sure event to all events. The probability of event A relative to W is called absolute probability.

Note that an impossible event has an absolute probability of zero, but, an absolute probability of zero does not imply that the event under consideration is impossible. An event that is as likely as the sure event does not need to be sure; it is called almost sure. An event A is called sure if its complement is impossible, i.e. $\pi(W \backslash A, \emptyset) = 1$.

In consequence of proposition 1 the rules of calculation that we know for probability measures are also valid for the absolute probabilities described by an RPM. So, the expected value of a $\mathcal{W}$–measurable function $f : W \to \mathbb{R}$ is

$$\mathrm{E}\left(f(z) \mid \pi\right) = \mathrm{E}\left(f(z) \mid \mu[\pi]\right).$$

Define $\mathcal{P}_{\mathcal{M}} := \{\pi \in \mathcal{P} \mid \pi(A, W) = 0 \;\Rightarrow\; \pi(\emptyset, A) = 1\}$.

$\mathcal{P}_{\mathcal{M}}$ is the set of RPMs for which events with probability zero are impossible. The next proposition shows that $\mathcal{P}_{\mathcal{M}}$ is an isomorphism to the set of probability measures on $\mathcal{W}$.

Proposition 2.A *Be $\mu \in \mathcal{M}$ and let $\pi[\mu] : \mathcal{W} \times \mathcal{W} \to \mathbb{R} \cup \{\infty\}$ be defined by*

$$\pi[\mu](A, B) = \begin{cases} \mu(A)/\mu(B) & \text{if } \mu(B) \neq 0 \\ \infty & \text{if } \mu(A) \neq \mu(B) = 0 \\ 1 & \text{if } \mu(A) = \mu(B) = 0. \end{cases}$$

Then $\pi[\mu] \in \mathcal{P}_{\mathcal{M}}$ and $\mu[\pi[\mu]] = \mu$.

Proposition 2.B *For each $\pi \in \mathcal{P}_{\mathcal{M}}$ it is true that $\pi[\mu[\pi]] = \pi$.*

The absolute probabilities, given by an RPM $\pi[\mu]$ are the same as the probabilities of these events expressed by μ. A probability measure μ can be replaced by an RPM $\pi[\mu]$, like a point expectation $z^e \in W$ can be expressed by a Dirac measure $\mu[z^e] \in \mathcal{M}$, defined by $\mu[z^e](\{z^e\}) = 1$.

5 Conditional Probabilities

The derivation of conditional probabilities is the heart of our considerations, and therefore, we want to give them an axiomatic foundation. Let the relative probabilities be measured by an RPM π, and denote the probabilities conditional to the occurrence of an event B by π_B.

Axiom 1 *Conditional relative probabilities can be represented by a relative probability measure, i.e.*

$$\pi_B \in \mathcal{P}.$$

Axiom 2 *After the occurrence of B the event $W \backslash B$ is impossible, i.e.*

$$\pi_B(W \backslash B, \emptyset) = 1.$$

Axiom 3 *Relative probabilities of subsets of B do not depend on the occurrence of B, i.e.*

$$A, C \subseteq B \quad \Rightarrow \quad \pi_B(A, C) = \pi(A, C).$$

These three axioms suffice to derive unique conditional probabilities whenever the condition is a possible event.

Proposition 3 *If $\pi \in \mathcal{P}$, $B \in \mathcal{W}$, and $\pi(\emptyset, B) = 0$, then there exists exactly one function π_B that obeys to the axioms 1 to 3 above. This function is characterized by*

$$\pi_B(A, C) = \pi(A \cap B, B \cap C) \qquad \forall A, C \in \mathcal{W}. \tag{1}$$

This justifies the following definition of a conditional RPM.

Definition 5 *Be $\pi \in \mathcal{P}$ and $B \in \mathcal{W}$. The* **conditional relative probability measure** *for π under condition B is the function $\pi_B : \mathcal{W}\times\mathcal{W} \to \mathbb{R}\cup\{\infty\}$, defined by (1).*

Interpretation $\pi_B(A, C) = k$ is interpreted: "Given the occurrence of event B, event A is k–times as likely as event C."

RPMs have the convenient property that conditional probabilities are uniquely defined for all possible conditions, even when the condition is an event with zero absolute probability. The conditional probability of event A given that event B occurres is the relative probability $\pi(A \cap B, B)$.

If conditional probabilities, as defined with a probability measure, are unique then the conditional probabilities derived from an RPM are the same.

Proposition 4.A *If $\pi \in \mathcal{P}$, $B \in \mathcal{W}$, and $\pi(B, W) > 0$, then*

$$\mu_B[\pi] = \mu[\pi_B].$$

Proposition 4.B *If $\mu \in \mathcal{M}$, $B \in \mathcal{W}$, and $\mu(B) > 0$, then*

$$\pi_B[\mu] = \pi[\mu_B].$$

Up to now, we have only shown the existence of RPMs contained in $\mathcal{P}_\mathcal{M}$. For those, a conditional RPM exists if and only if the conditional probability measure of the according probability measure is unique. The next proposition will demonstrate the existence of other RPMs.

Be $$\mathcal{P}^0[A] := \left\{ \pi \in \mathcal{P} \;\middle|\; \begin{array}{ll} \pi(\emptyset, B) = 0 & \text{falls } A \cap B \neq \emptyset \\ \pi(\emptyset, B) = 1 & \text{falls } A \cap B = \emptyset \end{array} \right\}.$$

$\mathcal{P}^0[A]$ is the set of RPMs, for which all events $B \subseteq A$ are possible and all events $B \subseteq W \backslash A$ are impossible.

$$\mathcal{P}^0[W] = \{\pi \in \mathcal{P} \mid \pi(\emptyset, B) = 0 \ \ \forall B \neq \emptyset\}$$

is the set of RPMs, for which $\emptyset$ is the only impossible event. The next proposition says that each probability measure can be completed to an RPM in $\mathcal{P}^0[W]$.

Proposition 5 *For each $\mu \in \mathcal{M}$ there exists a relative probability measure $\pi \in \mathcal{P}^0[W]$, such that $\mu[\pi] = \mu$.*

The proof of proposition 5 is given in the appendix. It constructs a hierarchy of probability measures, starting with μ. The succeeding measures are describing conditional probabilities for conditions that are null sets for all preceding measures. This procedure is closely related to the construction of a system of lexicographic probabilities as introduced by Blume, Brandenburger and Dekel (1991) for finite sets W, but also to the construction of a conditional probability space in Rényi (1976 [1956]).

The subset $\mathcal{P}^0[W]$ is isomorphic to the systems of conditional probabilities in the definition by Myerson (1986) and to the systems of lexicographic probabilities defined by Blume, Brandenburger and Dekel (1991). A RPM π is an isomorphism to the conditional probability measure $\mathbf{P} : \mathcal{W} \times \mathcal{B} \to [0,1]$ in the notion of Rényi (1976 [1955, 1956]) where $\mathcal{B}$ is the subset of all possible events,

$$\mathcal{B} := \{A \in \mathcal{W} \mid \pi(\emptyset, A) = 0\}.$$

Proposition 5 shows that the requirement $\pi(\emptyset, B) = 0$ is much weaker than $\mu[\pi](B) = \pi(B, W) > 0$. It directly follows from propositions 3 and 5 that if priors are modeled as an RPM $\pi \in \mathcal{P}^0[W]$ then posteriors are an RPM for all possible events. This feature is an advantage to modeling priors as conditional probability systems, for which posteriors are ordinary probability measures, so that there are two very distinct mathematical objects used to describe beliefs.

If an event A is impossible then $\pi_A = \pi_\emptyset$. We have seen above that $\pi_\emptyset$ violates axiom 7 of the definition of an RPM. Relative probabilities conditional on an impossible event are always one, which makes it difficult to interpret them.

The construction of conditional RPMs is a mathematical operation that knots an RPM with an event. The sure event W is neutral in this operation, since

$$\pi_W = \pi \qquad \forall \pi \in \mathcal{P}(\mathcal{W}).$$

An RPM that is neutral with regard to this operation does not exist in $\mathcal{P}$. As we have seen above, $\pi_A \in \mathcal{P}(\mathcal{W})$ requires $\pi(\emptyset, A) = 0$. Hence, the set $\mathcal{P}$ is not closed with regard to the operation π_A for all $A \in \mathcal{W}$.

Now, instead of $\mathcal{P}$, consider

$$\bar{\mathcal{P}} := \mathcal{P} \cup \{\pi_\emptyset\}.$$

Proposition 6 *For all $A \in \mathcal{W}$ it is true that*

$$\pi \in \bar{\mathcal{P}} \quad \Rightarrow \quad \pi_A \in \bar{\mathcal{P}} \qquad \text{and} \qquad \pi = \pi_\emptyset \quad \Rightarrow \quad \pi_A = \pi.$$

Proposition 6 shows that $\bar{\mathcal{P}}$ is closed with regard to the operation π_A and $\pi_\emptyset$ is the neutral element.

The interpretation of possible and impossible events suggests that after the occurrence of event B, any event $C \subseteq W \backslash B$ is impossible. An event $C \subseteq B$, that was possible under π, will also be possible under the condition that B occurres. This is implied by the next proposition.

Proposition 7 *If $A, B \in \mathcal{W}$, $A \cap B \neq \emptyset$, and $\pi \in \mathcal{P}^0[A]$, then $\pi_B \in \mathcal{P}^0[A \cap B]$.*

An immediate corollary is

Corollary 1 *If $\pi \in \mathcal{P}^0[W]$, $B \in \mathcal{W}$, and $B \neq \emptyset$, then $\pi_B \in \mathcal{P}^0[B]$.*

Consider the following scenario: An agent forms beliefs about the probabilities of events and successively gets the information that events A_1, A_2, A_3, ... occurred. An individual that has no information about the occurrence of events should consider all events as possible. Let us therefore assume that her a–priori–beliefs are an RPM $\pi \in \mathcal{P}^0[W]$. If we assume that she revises her beliefs according to the three axioms above, her posterior beliefs after getting the information $A_1, \ldots, A_n$ will be π_B, with $B = \bigcap_{i=1}^n A_i$.

Her beliefs will then be contained in a set $\mathcal{P}^0[B]$, where B is the intersection of all proceeded information. Given a posterior π_B the contained information can be extracted by

$$B = W \backslash \bigcup A \,\Big|\, \pi_B(A, \emptyset) = 1.$$

This shows that information, stored in a posterior belief, can be recalled. However, this requires that priors are taken out of $\mathcal{P}^0[W]$, since otherwise there may be events that had been considered as impossible right from the start and without respective information.

6 Independent Probabilities

Remember, two events A and B are called independent of each other, if $\mu(A \cap B) = \mu(A)\,\mu(B)$.

What is meant by the phrase "independent events"? An event is a set of states of the world. The phrase could be interpreted without any reference to probabilities: Let us define events $A := [a, b]$ and $B := [b, c]$. Here, A and B are dependent in the sense that a change in the parameters defining one event may have an impact on the other. Let $A := [a, b]$ and $B := [c, d]$. Here, A and B are independent of each other in the sense that a change in the defining parameters of one event does not influence the other.

In probability theory "independence" of A and B means
1. "the probability of event A is independent of the occurrence of event B"
and
2. "the probability of event B is independent of the occurrence of event A."

Using probability measures statement 1 [2] can be made only if B [A] is an event with positive probability:
1. $\mu(A) = \mu_B(A)$,
2. $\mu(B) = \mu_A(B)$.
But, if $\mu(B) = 0$, then $\mu_B(A)$ is not uniquely defined, and the validity of statement 1 can not be decided. For an RPM conditional probabilities are always unique. This allows for a better definition of independence:

Definition 6 *Be $\pi \in \mathcal{P}$ und $B \in \mathcal{W}$. The* **relative probability of two events** $A, C \in \mathcal{W}$ **is independent of event** B *if*

$$\pi(A, C) = \pi_B(A, C).$$

Using definition 6, we can say that the **absolute probability of an event** $A \in \mathcal{W}$ **is independent of event** $B \in \mathcal{W}$ if

$$\pi(A, W) = \pi_B(A, W).$$

For two events with positive absolute probability, definition 3 of independent events is equivalent to saying that the absolute probability of one event is independent of the other.

Proposition 8 *Be $\pi \in \mathcal{P}$, $\mu = \mu[\pi]$, and $\mu(B) > 0$. The absolute probability of A is independent of B if and only if A and B are independent events in the sense of definition 3.*

If the absolute probability of A is independent of B then the absolute probability of B does not need to be independent of A.

Example: If $A \subseteq B$ and $\pi(A, B) = \pi(B, W) = 0$, then

$$\pi(A, W) = \pi_B(A, W) = \pi(A, B) = 0 \quad \text{and}$$

$$\pi(B, W) = 0 \neq \pi_A(B, W) = \pi(A, A) = 1.$$

However, there still is a symmetry implied by the independence definition, which is

$$\pi(A, C) = \pi_B(A, C) \quad \Leftrightarrow \quad \pi(C, A) = \pi_B(C, A).$$

Intuitively one would call a series of events pairwise independent of each other if the probability of every one event is independent of the occurrence of the others.

Definition 7 *Be $\pi \in \mathcal{P}$. Events $A_1, A_2, \ldots$ are pairwise independent of each other if $\pi(A_i, W) = \pi_{A_j}(A_i, W)$, for all i, j with $i \neq j$.*

Corollary 2 *Be $\pi \in \mathcal{P}$ and $\mu = \mu[\pi]$. If events A and B are independent according to definition 7 then they are independent according to definition 3.*

The corollary follows immediately from proposition 8. This narrows the definition of independent events down to its intuitive meaning.

7 Conclusion

The paper introduced a theory of relative probability measures (RPMs). We have shown that one can calculate with relative probabilities as intuition suggests. This makes RPMs easy to handle. An RPM allows for a clear distinction between possible and impossible, sure and almost sure events. Conditional relative probabilities are uniquely defined for all possible conditions, even if the condition is a null–set. Furthermore, RPMs allow for a definition of independence that captures the intuition of independent events much better than the traditional definition.

If interpreted as subjective beliefs, an RPM is able to store information in such a way that it can be recalled from the posteriors.

Acknowledgements: *I am indebted to Hartmut Stein and Horst Stenger for their helpful suggestions.*

Appendix: Proofs

Proof of Lemma 1

1. Axiom 4 implies $\pi(A, A) \in \mathbb{R}_{++}$. Then, axiom 3 implies $\pi(A, A)\, \pi(A, A) = \pi(A, A)$. Hence, $\pi(A, A) = 1$.

2. If $0 < \pi(A, B) < \infty$, axiom 4 implies $0 < \pi(B, A) < \infty$. From axiom 3 and part 1 of this lemma we conclude $\pi(A, B)\, \pi(B, A) = \pi(A, A) = 1$. Hence, $\pi(A, B) = 1/\pi(B, A)$.

3. $A_n \nearrow B$ and axiom 5 imply $\pi(B \backslash A_n, B) \searrow \pi(\emptyset, B)$. From axiom 2 and

part 1 of this lemma we know that $\pi(A_n, B) = \pi(B, B) - \pi(B \backslash A_n, B) + \pi(\emptyset, B)$. Hence, $\pi(A_n, B) \nearrow \pi(B, B) = 1$.

4. For $\pi(B, W) > 0$ parts 2 and 3 of this lemma imply $\pi(W, B) = 1/\pi(B, W) \neq \infty$. Using axioms 3 and 7 we get $\pi(\emptyset, B) = \pi(\emptyset, W)\, \pi(W, B) = 0$.

5. If e.g. $\pi(A, B) = \infty$ and $\pi(B, C) \neq 0$ then axiom 4 implies $\pi(B, A) = 0$ and $\pi(C, B) \neq \infty$. From axiom 3 we get $\pi(C, A) = \pi(C, B)\, \pi(B, A) = 0$ and hence, $\pi(A, C) = \infty$ by axiom 4.

QED

Proof of proposition 1 Because of axiom 7 we have $\mu[\pi](\emptyset) = \pi(\emptyset, W) = 0$. Lemma 1, parts 1 and 3 imply $\mu[\pi](W) = \pi(W, W) = 1$ and $\mu(A) = \pi(A, W) \in [0, 1] \;\; \forall A \in \mathcal{W}$. Now, axiom 2 implies

$$\mu[\pi](A) + \mu[\pi](B) = \pi(A, W) + \pi(B, W)$$

$$= \pi(A \cup B, W) - \pi(A \cap B, W) = \mu[\pi](A \cup B) - \mu[\pi](A \cap B).$$

Finally, axioms 5 and 7 imply that for each sequence of events $\{A_n\}$, with $A_n \searrow \emptyset$, $\mu[\pi](A_n) = \pi(A_n, W) \searrow 0$. QED

Proof of proposition 2.A First, we show that $\pi[\mu]$ obeys to the seven axioms defining an RPM. Axiom 1 holds trivially.

2. If $\pi[\mu](A \cup B, C) \in \mathbb{R}_+$, we either have $\mu(A \cup B) = \mu(C) = 0$ or $\mu(C) > 0$. In the first case $\mu(A) = \mu(B) = \mu(A \cap B) = 0$ and therefore

$$\pi[\mu](A, C) + \pi[\mu](B, C) = 2 = \pi[\mu](A \cup B, C) + \pi[\mu](A \cap B, C).$$

In the second case

$$\pi[\mu](A, C) + \pi[\mu](B, C) = \frac{\mu(A) + \mu(B)}{\mu(C)} = \frac{\mu(A \cup B) + \mu(A \cap B)}{\mu(C)}$$

$$= \pi[\mu](A \cup B, C) + \pi[\mu](A \cap B, C).$$

3. If $\pi[\mu](A,B) \in \mathbb{R}_+$ and $\pi[\mu](B,C) \in \mathbb{R}_+$, we have either $\mu(A) = \mu(B) = \mu(C) = 0$ or $\mu(A) = \mu(B) = 0 < \mu(C)$ or $\mu(B) > 0 < \mu(C)$. In the first case $\pi[\mu](A,B)\,\pi[\mu](B,C) = 1 = \pi[\mu](A,C)$. In the second case

$$\pi[\mu](A,B)\,\pi[\mu](B,C) = 0 = \pi[\mu](A,C).$$

In case three

$$\pi[\mu](A,B)\,\pi[\mu](B,C) = \frac{\mu(A)}{\mu(B)}\frac{\mu(B)}{\mu(C)} = \pi[\mu](A,C).$$

4. $\pi[\mu](A,B) = 0 \;\Leftrightarrow\; \mu(A) = 0 \;\wedge\; \mu(B) > 0 \;\Leftrightarrow\; \pi[\mu](B,A) = \infty.$

5. Let $A_n \searrow \emptyset$. If $\mu(B) > 0$, $\pi[\mu](A_n \cap B, B) = \mu(A_n \cap B)/\mu(B) \searrow 0$. If $\mu(B) = 0$, $\pi[\mu](A_n \cap B, B) = \pi[\mu](\emptyset, B) = 1 \quad \forall n$.

6. Obviously, $\pi[\mu](\emptyset, A) = 0[1]$ if $\mu(A) > [=]0$.

7. $\pi[\mu](\emptyset, W) = \mu(\emptyset)/\mu(W) = 0$.

By definition $\pi[\mu](A,B) = 1$ whenever $\mu(A) = \mu(B)$. Hence, $\pi[\mu] \in \mathcal{P}_{\mathcal{M}}$. Furthermore,

$$\mu[\pi[\mu]](A) = \pi[\mu](A,W) = \mu(A)/\mu(W) = \mu(A).$$

QED

Proof of proposition 2.B By definition

$$\pi[\mu[\pi]](A,B) = \begin{cases} \pi(A,W)/\pi(B,W) & \text{if } \pi(B,W) \neq 0 \\ \infty & \text{if } \pi(A,W) \neq \pi(B,W) = 0 \\ 1 & \text{if } \pi(A,W) = \pi(B,W) = 0, \end{cases}$$

If $\pi(B,W) \neq 0$, axiom 3 and lemma 1, parts 2 and 3, imply

$$\pi[\mu[\pi]](A,B) = \pi(A,W)\,\pi(W,B) = \pi(A,B).$$

If $\pi(A,W) \neq \pi(B,W) = 0$, $\pi[\mu[\pi]](A,B) = \infty$. Suppose $\pi(A,B) \neq \infty$. Then axiom 3 would imply $\pi(A,W) = \pi(A,B)\,\pi(B,W) = 0$.

If $\pi(A,W) = \pi(B,W) = 0$, $\pi[\mu[\pi]](A,B) = 1$. Since $\pi \in \mathcal{P}_{\mathcal{M}}$, axiom 3 and lemma 1.2 imply $\pi(A,B) = \pi(A,\emptyset)\,\pi(\emptyset,B) = 1$. QED

Proof of proposition 3 It is straightforward to verify that the function π_B as defined by (1) obeys to axioms 1 – 6 in definition 4. Axiom 7 holds if and only if $\pi(\emptyset,B) = 0$. Then, it is obvious that axioms A1 – A3 hold for π_B. It remains to show that this is the only function for which axioms A1 – A3 hold.

Be π_B a function that obeys to axioms A1 – A3. Then $\pi_B(\emptyset, C\backslash B) = 1$ for all $C \in \mathcal{W}$, and

$$\pi_B(B\cap C, C) = \pi_B(C,C) - \pi_B(C\backslash B, C) + \pi_B(\emptyset, C)$$

$$= 1 - \pi_B(\emptyset, C\backslash B)\,\pi_B(C\backslash B, C) + \pi_B(\emptyset, C) = 1.$$

Be $A \in \mathcal{W}$. If $\pi_B(A,C) \in \mathbb{R}_+$ then

$$\pi_B(A,C) = \pi_B(A\cap B, C) + \pi_B(A\backslash B, C) - \pi_B(\emptyset, C)$$

$$= \pi_B(A\cap B, C)\,\pi_B(C, B\cap C) + \pi_B(\emptyset, A\backslash B)\,\pi_B(A\backslash B, C) - \pi_B(\emptyset, C)$$

$$= \pi_B(A\cap B, B\cap C) = \pi(A\cap B, B\cap C).$$

If $\pi_B(A,C)=\infty$ then $\pi_B(C,A)=0=\pi(C\cap B, B\cap A)$. Hence, $\pi(A\cap B, B\cap C)=\infty$. This shows that π_B is characterized by (1). QED

Proof of proposition 4.A If $\pi(B,W)>0$ then $\mu[\pi](B)>0$. Hence, $\mu_B[\pi]$ is unique, and for all $A\in\mathcal{W}$

$$\mu_B[\pi](A)=\frac{\mu[\pi](A\cap B)}{\mu[\pi](B)}=\frac{\pi(A\cap B,W)}{\pi(B,W)}=\pi(A\cap B,B).$$

Lemma 1.4 and proposition 3 imply that $\pi_B\in\mathcal{P}$. Therefore, $\mu[\pi_B]$ is defined, and

$$\mu[\pi_B](A)=\pi_B(A,W)=\pi(A\cap B,B).$$

Proof of proposition 4.B For all $A,B,C\in\mathcal{W}$ it is true that

$$\pi_B[\mu](A,C)=\pi[\mu](A\cap B,C\cap B)$$

$$=\begin{cases}\mu(A\cap B)/\mu(C\cap B) & \text{if } \mu(C\cap B)\neq 0\\ \infty & \text{if } \mu(A\cap B)\neq\mu(C\cap B)=0\\ 1 & \text{if } \mu(A\cap B)=\mu(C\cap B)=0.\end{cases}$$

If $\mu(B)>0$ then μ_B is unique and

$$\pi_B[\mu](A,C)=\begin{cases}\mu_B(A)/\mu_B(C) & \text{if } \mu_B(C)\neq 0\\ \infty & \text{if } \mu_B(A)\neq\mu_B(C)=0\\ 1 & \text{if } \mu_B(A)=\mu_B(C)=0\end{cases}$$

$$=\pi[\mu_B](A,C).$$

QED

Proof of propostion 5 Be $\mu \in \mathcal{M}$ and "$<$" a relation that well–orders[2] $\mathcal{W}$ such that $W < A < \emptyset$ for all $A \in \mathcal{W}\backslash\{\emptyset, W\}$. First, we construct two functions $\kappa : \mathcal{W} \to \Omega$ and $m : \Omega \to \mathcal{M}$, where Ω is the class of ordinals.

Set $m[0] := \mu$ and $\kappa(A) := 0$ for all A with $\mu(A) > 0$. Define

$$\mathcal{N}(\mu) := \{A \in \mathcal{W} \mid \mu(A) = 0\}.$$

Be $\mathcal{N}^1 := \mathcal{N}(\mu)$ and N^1 the first element of $\mathcal{N}^1$ with respect to the well–order "$<$". Set $\kappa(N^1) := 1$ and choose a probability measure $m[1]$ with $m[1](N^1) = 1$. Now, set

$$\kappa(A) := 1 \text{ for all } A \in \mathcal{N}^1, \text{ with } m[1](A) > 0.$$

In general, for any ordinal $\omega > 0$ set

$$\mathcal{N}^\omega := \bigcap_{k<\omega} \mathcal{N}(m[k]), \qquad \text{where } \mathcal{N}^0 := \mathcal{W},$$

$$N^\omega := \min_{<} \mathcal{N}^\omega,$$

$$m[\omega] \in \{\mu \in \mathcal{M} \mid \mu(N^\omega) = 1\} \text{ arbitrary, and}$$

$$\kappa(A) = \omega \quad \forall A \in \mathcal{N}^\omega, \text{ with } m[\omega](A) > 0.$$

[2] A well ordering relation on $\mathcal{W}$ is a binary relation with the following properties:

1. For all $A, B \in \mathcal{W}$, with $A \not\equiv B$, either $A < B$ or $B < A$.
2. For all $A, B, C \in \mathcal{W}$, $A < B \wedge B < C$ implies $A < C$.
3. For each $\mathcal{A} \subseteq \mathcal{W}$ there exists an $A \in \mathcal{A}$, with $A < B \quad \forall B \in \mathcal{A}\backslash\{A\}$.

The existence of a relation that well–orders $\mathcal{W}$ is implied by the Well–Ordering Theorem of Cantor und Zermelo (see Levy, 1979).

This definitory process is pursued until an ordinal k^* is reached, for which $\mathcal{N}^{k^*} = \{\emptyset\}$. Now, set $\kappa(\emptyset) := k^*$ and stop the process. Since $\mathcal{N}^k \subset \mathcal{N}^{k-1}$, the process stops at the latest with the ordinal $|\mathcal{W}|$, i.e. $k^* \leq |\mathcal{W}|$.

By construction of κ for all $A, B \in \mathcal{W}$,

$$\kappa(A) \geq \kappa(A \cup B) \tag{2}$$

$$m[\kappa(A)](B) > 0 \quad \Rightarrow \quad \kappa(A) \geq \kappa(B). \tag{3}$$

$$m[\kappa(A)](B) = 0 \quad \Rightarrow \quad \kappa(A) \neq \kappa(B). \tag{4}$$

Be $\quad \pi(A, B) := \begin{cases} \frac{m[\kappa(B)](A)}{m[\kappa(B)](B)} & \text{if } \kappa(A) \geq \kappa(B) \ \wedge \ B \neq \emptyset \\ \infty & \text{if } \kappa(A) < \kappa(B) \\ 1 & \text{if } A = B = \emptyset. \end{cases}$

Then 1. $\pi : \mathcal{W} \times \mathcal{W} \rightarrow \mathbb{R}_+ \cup \{\infty\}$,

2. If $\pi(A \cup B, C) \neq \infty$, then $\kappa(A \cup B) \geq \kappa(C)$. From (2) we conclude $\kappa(A) \geq \kappa(C) \ \wedge \ \kappa(B) \geq \kappa(C) \ \wedge \ \kappa(A \cap B) \geq \kappa(C)$.

If $A \cup B = C = \emptyset$, then

$$\pi(A, C) + \pi(B, C) = 2 = \pi(A \cup B, C) + \pi(A \cap B, C).$$

If $C \neq \emptyset$, then

$$\pi(A, C) + \pi(B, C)$$

$$= \frac{m[\kappa(C)](A) + m[\kappa(C)](B)}{m[\kappa(C)](C)} = \frac{m[\kappa(C)](A \cup B) + m[\kappa(C)](A \cap B)}{m[\kappa(C)](C)}$$

$$= \pi(A \cup B, C) + \pi(A \cap B, C).$$

3. If $\pi(A,B) \neq \infty$ and $\pi(B,C) \neq \infty$, then $\kappa(A) \geq \kappa(B) \geq \kappa(C)$.

If $A = B = C = \emptyset$, then $\pi(A,B)\,\pi(B,C) = 1 = \pi(A,C)$.

If $A = B = \emptyset \neq C$, then $\pi(A,B)\,\pi(B,C) = 0 = \pi(A,C)$.

If $B \neq \emptyset \neq C$ and $\kappa(B) > \kappa(C)$, then $\pi(B,C) = \pi(A,C) = 0$, because of (3)
.

If $B \neq \emptyset \neq C$ and $\kappa(B) = \kappa(C)$, then

$$\pi(A,B)\,\pi(B,C) = \frac{m[\kappa(B)](A)}{m[\kappa(B)](B)}\frac{m[\kappa(C)](B)}{m[\kappa(C)](C)} = \frac{m[\kappa(B)](A)}{m[\kappa(C)](C)} = \pi(A,C).$$

4. Because of (4)

$$\pi(A,B) = 0 \;\Leftrightarrow\; \kappa(A) > \kappa(B) \;\Leftrightarrow\; \pi(B,A) = \infty.$$

5. Be $\{A_n\}$ a sequence of events with $A_n \searrow \emptyset$. Because of (2) $\kappa(A_n \cap B) \geq \kappa(B)$. Hence, for $B \neq \emptyset$

$$\pi(A_n \cap B, B) = \frac{m[\kappa(B)](A_n \cap B)}{m[\kappa(B)](B)} \searrow 0 = \pi(\emptyset, B).$$

For $B = \emptyset$ we get $\pi(A_n \cap B, B) = 1 = \pi(\emptyset, B)$.

6. $\pi(\emptyset, A) = \begin{cases} \frac{m[\kappa(A)](\emptyset)}{m[\kappa(A)](A)} = 0 & \text{if } A \neq \emptyset \\ 1 & \text{if } A = \emptyset. \end{cases}$

7. $\pi(\emptyset, W) = \mu(\emptyset)/\mu(W) = 0$.

This shows that $\pi \in \mathcal{P}$.

Since $m[\kappa(B)](B) > 0\ \forall B \neq \emptyset$, $\pi(\emptyset, B) = m[k](\emptyset)/m[k](B) = 0$. Hence, $\pi \in \mathcal{P}^0[W]$. Since $\kappa(W) = 0$, $\mu[\pi](A) = \pi(A,W) = \mu(A)/\mu(W) = \mu(A)$ for all $A \in \mathcal{W}$. QED

Proof of proposition 6 Be $\pi \in \bar{\mathcal{P}}$. Proposition 3 implies that $\pi_A \in \mathcal{P}$ if $\pi(\emptyset, A) = 0$. If $\pi(\emptyset, A) = 1$, then $\pi_A(\emptyset, B) = \pi(\emptyset, A \cap B) = 1$ for all $B \in \mathcal{W}$. Thus $\pi_A = \pi_\emptyset \in \bar{\mathcal{P}}$.

If $\pi = \pi_\emptyset$, then $\pi_A(\emptyset, B) = \pi(\emptyset, A \cap B) = 1 = \pi(\emptyset, B) \quad \forall A, B \in \mathcal{W}$. Hence, $\pi_A = \pi$. QED

Proof of proposition 7 Be $A \cap B \neq \emptyset$ and $\pi \in \mathcal{P}^0[A]$. Then $\pi(\emptyset, B) = 0$. Thus, proposition 3 implies $\pi_B \in \mathcal{P}$. Be $C \in \mathcal{W}$. If $A \cap B \cap C \neq \emptyset$ then $\pi_B(\emptyset, C) = \pi(\emptyset, B \cap C) = 0$. If $A \cap B \cap C = \emptyset$ then $\pi_B(\emptyset, C) = \pi(\emptyset, B \cap C) = 1$. Therefore, $\pi \in \mathcal{P}^0[A \cap B]$. QED

Proof of proposition 8 Be $\pi \in \mathcal{P}$, $\mu = \mu[\pi]$, and $\mu(B) > 0$. Using proposition 1, Lemma 1.2, axiom 3 in definition 4, and equation (1), we get

$$\mu(A \cap B) = \mu(A)\, \mu(B) \quad \Leftrightarrow \quad \pi(A \cap B, W) = \pi(A, W)\, \pi(B, W)$$

$$\Leftrightarrow \quad \pi(A \cap B, W)\, \pi((W, B) = \pi(A \cap B, B) = \pi_B(A, W) = \pi(A, W).$$

QED

References

Aumann, Robert J. (1995) "Backward Induction and Common Knowledge of Rationality", *Games and Economic Behavior* 8, p. 6–19.

Battigalli, Pierpaolo, and Pietro Veronese (1996) "A Note on Stochastic Independence without Savage–Null Events", *Journal of Economic Theory* 70, p. 235-248.

Binmore, Kenneth G. (1992) "Foundations of Game Theory" in: J.–J. Laffont (ed.): *Advances in Economic Theory: Sixth World Congress*, Vol. I, Cambridge, UK: Cambridge Univ. Press, p. 1–31.

Blume, Lawrence, Adam Brandenburger und Eddie Dekel (1991) "Lexicographic Probabilities and Choice under Uncertainty", *Econometrica* 59, p. 61–79.

Heinemann, Frank (1995a) "Closing Economic Models by Information Assumptions", Discussion Paper 518-95, Insitut für Volkswirtschaftslehre und Statistik, Universität Mannheim.

Heinemann, Frank (1995b) *Rationalisierbare Erwartungen, Eine entscheidungstheoretische Fundierung ökonomischer und spieltheoretischer Gleichgewichtskonzepte*, Heidelberg: Physica.

Kohlberg, E., und P. Reny (1992) "An Interpretation of Consistent Assessments", mimeo.

Kolmogorov, A. N. (1933) Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin: Springer.

Levy, Azriel (1979) *Basic Set Theory*, Berlin: Springer.

McLennan, A. (1989) "The Space of Conditional Systems is a Ball", *International Journal of Game Theory* 18, p. 141–174.

Myerson, Roger B. (1986) "Multistage Games with Communication", *Econometrica* 54, p. 323–358.

Myerson, Roger B. (1991) *Game Theory: Analysis of Conflict*, Cambridge, Mass.: Harvard Univ. Press.

Reny, Philip J. (1995) "Rational Behavior in Extensive–Form Games", *Canadian Journal of Economics* 28, p. 1–16.

Rényi, Alfréd (1976 [1955]) “On a New Axiomatic Theory of Probability”, in: Pál Turán (Hrsg.), *Selected Papers of Alfréd Rényi*, Vol. 1, 1976, Budapest: Akadémiai Kiadó, S. 457–505. Originally published in *Acta Math. Acad. Sci. Hung.* 6 (1955), p. 285–335.

Rényi, Alfréd (1976 [1956]) “On Conditional Probability Spaces Generated by a Dimensionally Ordered Set of Measures”, in: Pál Turán (Hrsg.), *Selected Papers of Alfréd Rényi*, Vol. 1, 1976, Budapest: Akadémiai Kiadó, S. 547–557. Originally published in *UWCyr1214.5ptTeoriya Veroyatnostme£ i prim.* 1 (1956), p. 61–71.

Savage, Leonard J. (1972 [1954]) *The Foundations of Statistics*, 2^{nd} revised edition 1972, New York: Dover Publications. First edition 1954, John Wiley & Sons.

Stahl, Dale O. II (1995) “Lexicographic Rationality and Iterated Admissibility”, *Economics Letters* 47, p. 155–159.

Swinkels, Jeroen M. (1994) “Independence for Conditional Probability Systems”, Discussion Paper No. 1076, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston, Illinois.