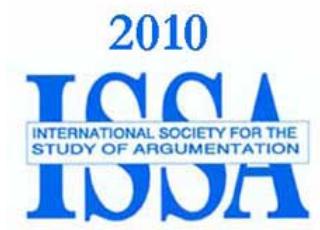# ISSA Proceedings 2010 – Probabilistic Arguments In The Epistemological Approach To Argumentation

1. Introduction: the Epistemological Approach to Argumentation and Probabilistic Arguments

In this paper I present a proposal on how to conceptualise and handle probabilistic arguments in an epistemological approach to argumentation. The epistemological approach to argumentation is an approach which aims at rationally convincing addressees or, more precisely, which takes knowledge or justified belief of an addressee to be the standard output of argumentation (Biro 1987, p. 69; Biro & Siegel 1992, pp. 92; 96; Siegel & Biro 1997, pp. 278; 286; Lumer 1990, pp. 43 f.; 1991, p. 100; 2005b, pp. 219-220; Goldman 2003, p. 58).[i] Therefore, this approach develops criteria for valid and adequate arguments whose observance leads, or at least is intended to lead, to the production of that output: justified belief. The general way in which this goal is achieved is by guiding the addressee through a process of recognising the truth or acceptability of the argument's thesis. An ordered sequence of judgements, i.e. the reasons, is presented to the addressee whose truths, according to a primary or secondary criterion of truth or acceptability, imply the truth or acceptability of the thesis and which are chosen in such a way that the addressee can immediately check whether they are true or acceptable (Lumer 1990, pp. 44-51; 2005b, pp. 221-224).

In an epistemological approach to argumentation, different types of arguments can be distinguished according to the respective epistemological principles on which they are based. There are e.g. *deductive arguments* based on deductive logic; there are *practical arguments* based on rational decision theory and its additions like game theory or philosophical theories of practical rationality; there are *empirical-theoretic arguments* for empirical laws about theoretical entities, which are based on criteria for good empirical theories; there are *probabilistic arguments* for probability judgements, which are based on probability theory; etc. To be constructively helpful, an epistemological theory of argumentation should

not only develop a general definition of 'good argument' but also elaborate precise criteria for such special types of argument. Such criteria have e.g. been proposed within the epistemological approach for deductive arguments (Feldman <1993> 1999, pp. 61-80; 94-100; Lumer 1990, pp. 180-209) or for practical arguments (Feldman 1999, pp. 351-354; 420; Lumer 1990, pp. 319-433). For the realm of probabilistic arguments, criteria for certain subtypes have been developed: criteria for *genesis of knowledge arguments* (which try to show that the thesis has been correctly verified by someone), which include *arguments from testimony* and *from authority* (Feldman 1999, pp. 216-232; 418; Goldman 1999, pp. 103-130; Lumer 1990, pp. 246-260), and for *interpretative arguments* (which try to establish the causes of known facts and circumstantial evidence through inference to the best explanation based on Bayes's Theorem) (Lumer 1990, pp. 221-246).**[ii]** A general theory of probabilistic argumentation which provides exact criteria for the validity and adequacy of these arguments is so far lacking, however. Such a theory will be proposed in the following, starting from the epistemological approach to argumentation.

"*Probabilistic argument*" here always refers to: an argument with a probability judgement as its thesis. Applying the epistemic approach to such arguments presupposes primary or secondary criteria of the truth or acceptability of probability judgements. Of course, such criteria should be provided by probability theory. But, although there is a rather broad consensus in probability theory about the calculus to be used, there is significant divergence about the interpretation and conceptualisation of probability, which would also lead to different conceptualisations of probabilistic arguments. So we first have to go some way into the philosophical debate on the best concept of probability.

2. Philosophical Concepts of Probability – A Case for Probability as Rational Approximation to Truth

Philosophical theories of probability come in two main groups: first, objective or realistic theories, which maintain that probability is an objective, real feature of the world, and, second, subjective, epistemic, or cognitivist theories, which maintain that probability is essentially an epistemic or belief phenomenon, due to our limited knowledge. The two main realistic approaches are, first, *relative frequency theories*, according to which probabilities are identical to actual relative frequencies (Venn <1866> 2006) or to limiting relative frequencies in a hypothetical infinite row of trials (Reichenbach <1935> 1949; von Mises <1928>

1981), and, second, *propensity theories*, according to which probabilities are identical to a quantitative disposition in an object or in a type of system to produce a certain result or results with a certain relative frequency (Gillies 2000; Hacking 1965; Mellor 2005; Miller 1994; Popper 1959). Propensity theories have been developed to explain single-event stochastic processes like radioactive decay of single atoms, whereas frequency theories seem to capture particularly well probabilities derived by statistical inferences.

At first appearance, only realistic or objective theories seem to be appropriate to provide what an epistemological approach to argumentation needs, namely objective criteria for the truth of probability judgements. This impression, however, is due to an ambiguity of the word "objective". A judgement may be "objective", in a weak sense, of being cognitive, i.e. of being true or of being the result of an interpersonally verifiable process of applying clear criteria. And a judgement may be "objective", in a stronger sense, of being realistic, i.e. of describing a reality that is independent of any subjective attitude. Of course, only realistic theories of probability are objective in the strong sense; however an epistemological approach to probabilistic arguments needs objective criteria for the truth or acceptability of probability judgements only in the weak, cognitivist sense. This weak kind of objectivity can, however, also be provided by some epistemic theories of probability so that the objectivity requirement is no argument in favour of realistic theories of probability.

There are many well-known objections against every single realistic theory of probability. (In theories of actual frequencies, e.g. the result of a series of experiments may strongly diverge from the true probability – think of a die rolled only three times in its life and always showing "6" (Hájek <2002> 2009). In theories of limiting relative frequency real infinite series are impossible and soon lead to radical changes of the experimental situation – what a die will look like after having been rolled a billion times? –, whereas hypothetical infinite series have left empiricism behind (ibid.). Propensity theories share many problems of frequentism; in addition, propensities are causalist and hence asymmetric, whereas probabilities in a certain sense may be "inverted" – Bayes's Theorem e.g. implies such an inversion of conditional probabilities: $P(a/b) = (P(b/a) \cdot P(a)) / P(b)$ –; it may make sense to say that affluent people have a propensity to vote for conservative parties, whereas it makes little sense to say that votes for conservative parties have a propensity to come from affluent people (Humphreys

1985).). I want to stress here, however, only two general objections. The first is ontological. Of course, there are relative frequencies and these provide us with information about probabilities, and there are qualitative structures of a system underlying these relative frequencies. But what are the realistically conceived frequentist or propensity probabilities of a single event? Either the die ends up with "6" on top or it does not; and if it does this was probably determined by laws of nature. We are speaking of probabilities in such cases only because we do not know the result beforehand; we try to approach truth as much as possible before the event by speaking of probabilities. Afterwards our probabilities even change, e.g. to the probability 1 for "6". At least probabilities of single events are epistemic probabilities; and for the objective fact of relative frequencies we have precisely the notion of 'relative frequency', which is different from 'probability'. Probabilities are only an epistemic substitute in case of incomplete knowledge.

The other general objection to realist theories of probabilities is particularly relevant to our endeavour to develop an epistemic theory of probabilistic arguments. There are many epistemic uses of probabilities which do not try to capture real tendencies in the world. This holds in particular when we try to find out backward information, like the probable cause of a known fact – e.g. 'the dinosaurs probably (with a probability of 90%) became extinct as a consequence of a giant asteroid hitting the Earth' (cf. Hacking 2001, pp. 128-130) – or the probable meaning of a sentence or the fact indicated by a clue. Such probability statements do not speak of frequencies or propensities in the world but even try to fill our gaps of knowledge backwards. Hence these probabilities are quite obviously epistemic in nature.

What, then, about epistemic theories of probability? We have to dismiss rather quickly the traditional Laplacian *theory of a priori equiprobability of logical possibilities*, which has the immense disadvantage of not incorporating empirical information about relative frequencies,[iii] and the *theory of logical probabilities or inductive logic* (Carnap 1950; 1952). Problems of the latter theory, among others, are that its confirmation function is arbitrary or that, contrary to what the theory presupposes, (basic) evidences do not necessarily have the probability of 1. The major remaining approach then is *subjectivism* or *personalism* or *subjective Bayesianism*, which conceives probabilities in a personal or subjective way as rational degrees of belief.

Subjectivism conceives degrees of belief in a behaviouristic manner as something

revealed by preferences. In the most simple case a subjective probability p of an event e is equated with that value p for which it is true that the subject is indifferent about receiving some amount of money $p \bullet m$ for sure and a lottery by which the subject receives the complete amount m conditional on e. ($Ps(e)=p :=$ $p \bullet m \approx s <e; m; \neg e; 0>$). More complex systems make stronger presuppositions about preferences and measure probabilities as well as utilities (Eells 1982, pp. 9 f.). Behaviouristic conceptions of the degree of belief lead to well-known problems, e.g.: buying and selling prizes usually differ; the utility function of money is not linear, hence the utility of $p \bullet m$ is not identical to p times the utility of m. The general problem behind such difficulties is the behaviouristic approach, which has to find out too many interdependently acting subjective variables only on the basis of knowledge about the behavioural surface.

Let me now add to this a further and less well-known aspect of this problem, which is detrimental to the usual interpretation of subjectivism itself. The standard interpretation of subjectivist probability as degree of belief cannot, though we actually have to, distinguish between, first, a (less than certain) degree of belief or confidence (in a non-technical sense) and, second, a belief with a probabilistic content. That we have to distinguish these two things is obvious in situations where both phenomena are present. Someone has heard from an expert that the probability of some event e is p, or he has inferred this probability from his own frequency counts and hence believes that the probability of e is p. He is not sure about this probability, however, and has only a reduced degree q of belief in it – e.g.because he knows he has a bad memory or cannot recall clearly the value p or because he has some doubts about the expert's reliability. The difference between the two kinds of uncertainty is that the probability believed is part of the belief's content, i.e. the proposition believed in, whereas the degree of belief is outside this propositional content as it is something like the intensity of the propositional attitude. Instead of simply saying 'subject s believes that a' **[iv]** and thus taking 'belief' as a qualitative notion we can take it as a quantitative, functional notion: 's believes that a to the degree q (or with the confidence q)' and we can write this as: $Bs,a=q$. If the proposition believed in has a probabilistic content, as in the example just given, we can write this as: s believes with the confidence q that the probability of e is p: $Bs,(Pe=p)=q$. This differentiation, however, constitutes a problem for the usual subjectivist interpretation of probability. The probability now shows up already in the belief's content; what does this (inner) concept of 'probability' then mean? If the probability value and

the confidence value differ – as they are supposed to do in most cases – then this probability cannot be the degree of belief. At least it cannot be the degree of belief of that person at that time. The defender of subjective probabilities as degrees of belief may reply to this objection: but it can be the degree of belief of a different person or of the same person at a different time. From the several alternatives – e.g. the subject's earlier belief, the informant's belief or the belief of a rational subject – the latter seems to be the most plausible because this interpretation would be possible in any case and not only in a limited number of cases. However, this proposal faces serious problems too. First, the differentiation in believed probability and degree of belief seems to exist already just from the beginning even for very rational subjects who e.g. have determined some probability on the basis of a frequency count executed a second ago but, because of considering human fallibility, have only a confidence near to one. To what other degree of certainty shall the probability judgment refer to in this case? Second, rational subjects should be exactly the people who base their subjective probabilites on clear epistemic procedures, which are different from simply having a certainty impression (which, perhaps, might be interpreted as the degree of certainty). According to an at least slightly verificationist semantics, some of these procedures would make up the meaning or content of the resulting belief's *proposition*, so that (1) the probabilistic content would belong to the propositional content and would not make up the external degree of confidence and (2) it would have a meaning other than referring to a degree of belief. All this means we are still lacking an interpretation of 'probability'.

What then are probabilities? My proposal for answering this question is: In alethic terms, probabilities are *rational approximations to truth* under conditions of epistemic limitations. This is supposed to mean that the respective subject does not know whether the real value is 0 or 1, which are the only possible values; but his knowledge, though not sufficient to establish 0 or 1, indicates a value between these extremes, which may be nearer to 0 (or 1) to a given degree. Given these explanations, an explanation of probabilities in epistemic terms seems to be even more adequate. Therefore, instead of speaking of "approximation to truth" one might also say that probabilities, in justificatory terms, express a certain *tendency of evidence* for the two possibilities, which we can extract from our limited knowledge. If the probability of an event e is p the tendency of evidence for e is p, with $0 \leq p \leq 1$, whereas the tendency of evidence for e being false is $1-p$.[v] In practical terms, finally, probabilities are *degrees of rational reliance* that the

event in question will occur. They are values we ascribe to propositions for decisional purposes and which by maximising expected utility permit us to follow a strategy which, according to the laws of large numbers, in the long run will be the best among the strategies we can follow with our limited knowledge. These three aspects coincide because the epistemic aim is exactly to approach truth as closely as possible with the given information; and using these approximations in one's decisions in a decision-theoretic fashion implies making maximum and specific use of the information at hand.

3. Some Syntactical Features of Probabilities as Tendencies of Evidences
According to the explanations just given, probabilities as tendencies of evidences depend on a given corpus of knowledge; i.e. from different corpora of knowledge may result different degrees of probability: after having witnessed the rolling of a die the probability of showing "6" may increase from 1/6 to 1 (or decrease to 0). For nonetheless being able to be objective in the sense of being true, probability judgements have to express this kind of relativity by a respective variable that refers to the particular knowledge on which the probability is based.

Given the rationality and cognitivity of the probability striven for, the truth of a probability judgement should not depend on the identity of the believer but on the particular data corpus of which the person disposes. A different person with the same data corpus should, of course, assume the same rational probabilities. This means the knowledge variable of the probability concept should refer to data bases and not to persons (and moments). Of course, this does not exclude that the intended data base is denoted by a definite description that identifies the data with those at hand to a certain person at a certain time: 'Susan's data at that moment'. In ordinary language as well as in theoretic expositions the reference to the data base is rarely expressed explicitly; often it is simply identical to the speaker's data base at that very moment. (The probability relativised in this way, e.g. 'the probability of event e on Susan's data base is q', has to be distinguished from the subjective, or more precisely, from the *believed probability*, which can be expressed with our probability concept too: 'Susan believes that the probability of the event e (on her present data base) is p.') A further advantage of taking data corpora as the second variable of probabilities is that in this way things like 'scientific' or even 'natural probabilities' can easily be defined. A *scientific probability* would be one where the data base is the present scientific knowledge. And a *natural probability* of an event could be one where the data base is a

complete (true) description of the world's history before that event plus the (true) natural laws.

The other variable – or, in the case of conditional probabilities, the other two variables – of the probability concept refer to the things whose probability is expressed. Sometimes it is assumed that these relata are events, or more generally, states of affairs. This may be true in a realist approach to probability; however, in an epistemic approach the relata have to be what can be the content of knowlegde, i.e. propositions. To put it another way, the relata of epistemic probabilities have to be more fine-grained than events, namely propositions, because, though 'Peter's murderer has poisoned him' and 'Sara has poisoned Peter' could well denote the same event, the respective data base may not imply that Sara was Peter's murderer so that the probabilities of the two tokens may be different. And this is possible only if, given the identity of the event and of the data base, the tokens are propositions.

So, finally, the syntax of basic probability judgements is: 'The probability of the proposition a on the data base d is x' ($Pa,d=x$ or, if one prefers brackets: $P(a,d)=x$), and of conditional probability judgements it is: 'The probability of a given b on the data base d is x' ($P(a/b),d=x$ or $P((a/b),d)=x$). (Pa,d does not coincide with $P(a/d)$ because the d in the first instance is supposed to be true but in the second instance it is not. Nor does $P(a/b),d$ coincide with $P(a/d)$ because d does not need necessarily to imply b.)

4. Justifications of Probability Judgements: 1. Basic Probabilities

How can probability judgements be recognised in an epistemologically qualified way? In the realm of (more or less) certain knowledge we distinguish between basic or elementary cognition, in particular observation, on the one hand and derivative cognition proceeding by deductive inferences on the other. In the realm of probabilistic knowledge we can distinguish in a similar way between basic cognitions of probabilities via known relative frequencies (these cognitions do not rely on probabilistic premises, hence provide basic probability judgements) and derivative cognitions of probabilities by applications of the probability calculus, which already uses probabilities as inputs.

The basic form of probability cognitions of e.g. whether e or $\neg e$ via known relative frequencies, trivially, works as follows. It, first, presupposes that we have no better information about e, e.g. no definite information that e happened. It further

presupposes that we know some relative frequencies applying to *e*, i.e. relative frequencies of the form: 'The relative frequency of *E*s among *F*s is *x*', where *e* has the property *F* and perhaps the property *E*. Finally, it presupposes that if there are several such relative frequencies we can identify the one which is most specific about *e*, i.e. entails the most detailed description *F* of *e*. In a certain sense this specificity condition is a further special case of the condition that the data base does not contain any further information by means of which we can draw stronger conclusions about *e*. If all these presuppositions are fulfilled we can infer that the probability of *e* is *x*. (We may formalise these conditions as follows: "*RF(E/F)=x*" shall mean: the relative frequency of *E*s among *F*s is *x*; "*NBI*" shall mean: "no better information", i.e. the precedent information is the best in the respective data base about the proposition in question. With these abbreviations the conditions can be formalised as:

Foundation Principle:

$$P(e \mid RF(E/F){=}x \And f \And NBI),d = x),$$

for all *E, F, d*, *e*, *f* and *x* with $P(RF(E/F){=}x \And f \And NBI),d > 0$.

(This Foundation Principle is a reformulation of Hacking's Principle of Direct Probability (Hacking 1965; 2001, p. 137).

Note that this Foundation Principle does not presuppose any probabilistic information as an input of its use: relative frequencies are objective realities, which sometimes can be known with certainty; the same holds for "*f*", i.e. the fact that the possible event *e* has the quality *F*. Thus the Foundation Principle is really basic in the sense of newly introducing probabilities without already presupposing other probabilities.

The use of the Foundation Principle and hence the use of basic probabilities can be justified practically, i.e. as practically rational, on the basis of the laws of large numbers. If we do not dispose of certain information, probabilistic beliefs acquired via the Foundation Principle are the most informative condensation of our information about the event in question. If we use them via expected utility maximisation, of course, this cannot guarantee success in any single case but in the long run will provide better results than the use of any other way of handling uncertain information; as can be shown in comparisons with other decision strategies expected utility maximisation will lead to the highest utility. This

justification, however, does not say anything about the success of expected utility maximisation in any single case. So there may be decision situations where the large number presupposition does not hold – e.g. in decisions about life and death, where a fatal result implies simply that there will not be any further risky decision – and where expected utility maximisation may not be the best decision strategy. Hence, the just mentioned practical justification of the Foundation Principle proves the usefulness of employing probabilities calculated by means of this principle in many situations and justifies the use of the utility maximising strategy in many situations but it does not justify always weighting probabilities in decision situations according to the identity function, i.e. the probability x with the weight x.

Counting the magnitude of the population and the quantity of the positive cases is the safe way to establishing relative frequencies. This is costly, however, and not always possible. Therefore we need further ways to acquire information about relative frequencies. One less secure way is to try to remember single occurrences of the relative frequency in question and to count them. In addition, fortunately, mother nature has provided us with a not very reliable but at the same time not too bad sense of relative frequencies; on the basis of this we may consider past experiences and estimate in a holistic way their relative frequencies. This sense of relative frequencies can also lead to an uncertain degree of belief in a universal connection of two types of events. Another way to obtain information about relative frequencies, then, is to rate one's degree of certitude about such a connection and to take it as the relative frequency. This strategy may be called "propositionalisation of degrees of certitudes" because the degree of certitude, which is the intensity of the belief and hence not part of its content, is now made available as quantitative information within the beliefs proposition. This makes the quantitative information universally usable.

Propositionalisation of certitudes: $Ppf(RF(E/F)=y \mid Bs,("x(Fx®Ex))=y \& NBI),d=1$, for all s, E, F, y, d with *$P(Bs,("x(Fx®Ex))=y \& NBI) > 0$,*
where *Ppf* is a *prima facie* probability, which may be combined with other *prima facie* probabilities to obtain the final probability.

The final and the weakest way of acquiring information about relative frequencies presupposes that the data base contains absolutely no empirical information about the case in question. In such a situation we may establish relative frequencies in a Laplacian way by counting the logical possibilities.

The methods of establishing or estimating relative frequencies described so far scrutinise all the individuals of the population, which is often too expensive or even impossible. The range of these methods can be enormously extended if the scrutinised set can be considered as a (more or less) representative sample of a much bigger population so that the relative frequency established in the sample may be extrapolated as holding for the whole population. Statistics and considerations about projectability of properties tell us when and with which degree of confidence this can be done.

5. Justifications of Probability Judgements: 2. Derivative Probabilities
The other way to cognise probability judgements in an epistemologically qualified way is to calculate probabilities with the help of the probability calculus. Fortunately, this technical part of probability theory is much less controversial; a certain orthodoxy has been achieved. My task here is therefore only to remember some basic principles of this calculus. The basic axioms of the calculus are:
Normalcy: For all *a* and *d*: $0 \leq Pa,d \leq 1$.

Certainty: Certain propositions have the probability 1.

Additivity: If a and b are mutually exclusive then: $P(a\acute{U}b),d = Pa,d + Pb,d$, for all *a, b, d*.

Conditional probabilities: $P(a/b),d = (P(a\&b),d)/(Pb,d)$, for all *a, b, d* with $Pb,d>0$.

From these axioms follow theorems like:

Overlap: If *a* and *b* are not mutually exclusive then: $P(a\acute{U}b),d = Pa,d + Pb,d - P(a\&b),d$, for all *a, b, d*.

Complementarity: $P(\neg a),d = 1-P(a),d$, for all *a* and *d*.

Bayes's Theorem, extended: Let *h*1 to *hn* be mutually exclusive and exhaustive hypotheses, and *e* some relevant evidence, then:

$P(e/hi),d \bullet P(hi),d$

$P(hi/e),d = $ ¾¾¾¾¾¾¾¾¾¾¾

$j=1ånP(e/hj),d \bullet P(hj),d$

6. Rules for Derivative Probabilistic Arguments

As described in the introduction, according to the epistemological approach to argumentation, arguments should be able to guide an addressee in a process of recognising the acceptability of the argument's thesis. And they do this by presenting him reasons, i.e. judgements, which according to an epistemological primary or secondary criterion for the acceptability of the thesis imply this acceptability. The addressee may then check the truth of these reasons and of the implication relation and thus convince himself of the thesis's acceptability.

So a very simple probabilistic argument may look like this:
Thesis q: The probability of rolling a "1" or a "2" in the next cast is 1/3.

Indicator of argument: This holds because:

Reason r1: The additivity axiom of the probability calculus says that probabilities of mutually exclusive possibilities add up to the probability of the disjunctively combined event.

Reason r2: The probability of rolling a "1" (in the next cast) is 1/6.

Reason r3: The probability of rolling a "2" (in the next cast) is also 1/6.

Reason r4: The possibilities of rolling a "1" and of rolling a "2" are mutually exclusive.

Reason r5: 1/6 + 1/6 = 1/3.

Hence the thesis.

A formal version of this argument may be clearer:

Thesis q: $P(\text{"1"Ú"2"}), d_i = 1/3$ – with d_i referring to a particular data base, e.g. Peter's knowledge exactly at 12 noon (five seconds later Peter may already know e.g. that "1" is true, hence: $P(\text{"1"Ú"2"}), d_j = 1$).

Indicator of argument: Proof:

r1: Additivity: If *a* and *b* are mutually exclusive then: $P(a \text{Ú} b), d = Pa, d + Pb, d$, for all *a, b, d*.

r2: $P(\text{"1"}), d_i = 1/6$.

r3: P("2"),d_i=1/6.

r4: P("1"&"2"),d_i=0.

r5: 1/6 + 1/6 = 1/3.

Q.e.d.

In everyday life such explicit and extended arguments are virtually non-existent. But we may find abbreviated versions of them like this: "The probability of rolling a "1" or a "2" in the next cast is 1/3 because the probabilities of both these possibilities individually are 1/6; and because the two possibilities exclude each other their probabilities have to be added, which makes 1/3." So in this abridged version the reference to the data base is missing as well as the quote of the additivity axiom; and the mention of the mutual exclusiveness may be missing as well. Of course, for representing a valid argument the parts omitted in such an abridged argument must hold nonetheless and they must be reconstructable for an addressee. So we have to distinguish ideal, complete probabilistic arguments and non-ideal abridged versions of them whose validity is defined in terms of a corresponding ideal argument.

Following these indications, I have tried to provide a reasonably precise definition of 'valid derivative probabilistic argument' in two steps, by first defining what an *'ideal* valid derivative probabilistic argument' is and then giving the general definition.

x is an *ideal (argumentatively) valid derivative probabilistic argument*, iff *x* satisfies the conditions PA0 to PA3.

PA0.1: Domain of definition: x is a triple <r°,i,q>, consisting of

1. a set *r°* of judgements *r*1, *r*2, ..., *rn*,

2. an indicator *i* of argument, and

3. a judgement *q*.

*r*1, ..., *rn* are called the "reasons for *q*" and *q* is called "the thesis of *x*".

PA0.2: Structure of the argument:

PA0.2.1: Type of thesis: q is of the form: 'The probability of *a* (given *b*) on the data base *d* is *p*.' (*Pa,d=p* or *P(a/b),d=p*).

PA0.2.2: Kinds of reasons:

1. At least one of the reasons *r*1, …, *rn* is an axiom or theorem of the probability calculus, hence a general probabilistic judgement.

2. The singular probability judgements among the reasons all refer to the data base *d* (cf. PA0.2.1) or in part to *d* and the other part to a predecessor *dprior*, i.e. *d* without some evidence *e* (*dprior = d\e*).

PA1: Indicator of argument: i indicates that *x* is an argument, that *r*1, …, *rn* are the reasons and that *q* is the thesis of *x*. In addition, *i* can indicate that *x* is a probabilistic argument.

PA2: Guarantee of truth:

PA2.1: True premises: The judgements *ri* are true.

PA2.2: Inferential validity: The axioms and theorems of the probability calculus contained in *r*° and the other reasons perhaps contained in *r*° imply mathematically *q* – i.e. according to deductive and arithmetic rules.

PA2.3: Best evidence: d does not contain information that permits stronger conclusions about *a* (or, respectively, about the conditional probability *P(a/b),d*).

PA3: Adequacy in principle: x fulfils the standard function of arguments; i.e. *x* can guide a process of recognising the truth of *q*.

PA3.1: The reasons *r*1 to *rn* are well-ordered, i.e. as chains of equations and insertions of data in general formulas.

PA3.2: Apart from intermediate results, *r*° does not contain reasons that are superfluous for fulfilling the derivability condition PA2.2.

PA3.3: There is a subject *s* and a time *t* for which the following holds:

PA3.3.1: the subject *s* at the time *t* is linguistically competent, open-minded, discriminating and does not know a sufficiently strong justification for the thesis *q*;**[vi]**

PA3.3.2: d refers to *s'* data base at *t*; and

PA3.3.3: if at *t x* is presented to *s* and *s* closely follows this presentation this will make *s* justifiedly believe that the thesis *q* is acceptable; this process of cognition will work as follows: *s* will follow the chains of equations and insertions affirmed in *r°*, check their truth, thereby coming to a positive result.

Explanation regarding PA0.2.1 and PA0.2.2: The thesis of a probabilistic argument in the sense used here is a singular probability judgement, i.e. judgement which attributes a specific probability to a specific proposition. So general probability judgements, i.e. in particular theorems of the probability calculus, are not included for the simple reason that such theorems can be justified in deductive arguments, deriving them deductively from the axioms of the probability calculus – like any mathematical theorem. Such arguments, as opposed to probabilistic arguments, do not depend on the particular data base; their theses are general judgements quantifying over *any* data base *d* (cf. the examples given in sect. 5), they are not relative to a particular data base (as d_i in the example given at the beginning of this section). Only the dependence on a specific data base requires the particular conditions of probabilistic arguments such as the conditions 'best evidence' (PA2.3) or 'data base' (see below, PA5.5).

x is a *(argumentatively) valid derivative probabilistic argument*, iff *x* satisfies the conditions PA4.1 or PA4.2.

PA4.1: Ideal argument: x is an ideal valid derivative probabilistic argument, or

PA4.2: Abridged argument: x is not an ideal valid derivative probabilistic argument, but there is such an (ideal valid derivative probabilistic) argument *y* which to a certain extent is identical with *x* but for which the following holds:

1. The set of reasons *rx°* of *x* is a subset of the set of reasons *ry°* of *y* or of abridged versions of these reasons (cf. PA4.2.3).

2. The reasons perhaps missing in *rx°* are axioms or theorems of the probability calculus or they represent intermediate results; and the chain of equations is not interrupted by these omissions.

3. In the thesis *q* or in some of the probabilistic reasons of *x* the reference to the data base *d* may be omitted.

4. Condition PA3.3 holds analogously also for *x*.

Valid arguments are instruments for fulfilling a certain function, namely the function to lead to the cognition of the thesis; like all instruments they can fulfil their function only if they are used properly. In particular the valid argument must fit with the addressee's cognitive situation. In the following the adequacy conditions for an epistemically successful use of probabilistic arguments for rationally convincing are sketched. The most particular among these conditions is that the data base referred to in the argument has to be more or less identical to the data base of the addressee (PA5.5).

A valid probabilistic argument *x* is *adequate* for rationally convincing an addressee *h* (hearer) at *t* of the thesis (*q*) of *x* and for making him adopt the thesis' probability for himself iff condition PA5 holds:

PA5: Situational adequacy:

PA5.1: Rationality of the addressee: The addressee *h* (at *t*) is linguistically competent, open-minded, discriminating and does not have a sufficiently strong justification for the thesis *q*.

PA5.2: Argumentative knowledge (of the addressee): The addressee *h* at *t* knows at least implicitly the idea of the probability calculus and the mathematics used in *x*.

PA5.3: Explicitness: If *x* is not an ideal argument such that $r°$ does not contain all the reasons of the corresponding ideal argument the addressee *h* at *t* is able to add the most important of the missing reasons.

PA5.4: Acceptance of the reasons: The addressee *h* at *t* has recognised the truth of the reasons r_i of *x* and, in the case of non-ideal arguments, of its corresponding ideal or is able to recognise them immediately. And

PA5.5: Data base: The data base d_{ht} of *h* at *t* is identical to *d* or so near to *d* that the resulting probabilities regarding the reasons r_i and the thesis *q* remain unaltered.

The just defined probabilistic arguments are not special kinds of deductive arguments or reducible to them. One highlight of the present approach is to make the relativity of probabilistic arguments to specific data bases explicit, by

inserting a reference to the data base d, thus resolving the problems of logical non-monotonicity. As a consequence of this *explicit* relativity to the data base, the arguments can be and have to be (cf. PA2.2, inferential validity) deductively valid; in addition their reasons can be *true* – even the singular probability judgements among the premises. The problem of probabilistic arguments, i.e. to be only a substitute for stronger arguments in case of insufficient knowledge, which leads to non-monotonicity, however, cannot be eliminated entirely. Here it has been shifted to the pragmatic adequacy conditions, where PA5.5 requires to use an argument with a data base fitting to the addressee. Of course, the addressee may be convinced by an argument that refers to a different data base d_j that the probability of an event a on the data base d_j is p_j; however, if d_j is not the addressee's data base at the time being he will not adopt p_j as *his* probability. Deductive arguments do not contain any comparable restriction because they are not relative to the data base; for being rationally convincing the addressee has to be convinced of their premises, yes; but this is not yet a general dependency on the data base. Instead of being logically non-monotonic, probabilistic arguments as they are conceived here are "*pragmatically* non-monotonic" in the sense of getting pragmatically irrelevant when the data base does no longer fit to the addressee's changed data base. Further irreducible differences with respect to deductive arguments then are that this relativity to the data base also shows up in the adequacy in principle condition (PA3.3.2), that references to a data base are part of the structure of ideal probabilistic arguments (cf. PA0.2.1, PA0.2.2.2) and, finally, the best evidence requirement (PA2.3).

The just provided definitions show that it is possible to develop clear, reasoning-guiding and epistemologically justified criteria for probabilistic arguments, which do justice to requirements of objective validity as well to adaptation to the specific epistemic limits of the argument's addressees.**[vii]**

NOTES
[i] Proponents of the epistemological approach to argumentation are e.g. Mark Battersby, John Biro, Richard Feldman, Alvin Goldman, Christoph Lumer, Harvey Siegel and Mark Weinstein. An overview of this approach (including bibliography) is provided in: Lumer 2005a.
[ii] Several other forms of probabilistic arguments and fallacies have been analysed (e.g. Korb 2004; Hahn & Oaksford 2006; 2007), without however providing precise criteria for such arguments.

[iii] This dismissal as a general theory does not exclude that equiprobability settings play an important role in situations under complete uncertainty about frequentist probabilities.

[iv] Here and in the following I omit the time variable of 'belief'.

[v] The tendency of evidence should be distinguished from the degree or strength of evidence. We may have strong or weak evidence with the same tendency, i.e. for the same probability. We may e.g. have counted 30 black and 60 white balls before putting them into an urn and therefore have strong evidence that the probability of picking a white ball at random is 2/3; and, in a different setting, we may have picked (with replacement) nine balls from the urn, three of them being black and six of them being white, and because of this have the weaker evidence that the probability of picking a white ball at random is again 2/3.

[vi] That *s* is "linguistically competent" shall mean that she knows the semantics, syntax and expressions used in the argument; this includes knowledge about the probability concept and the parts of the probability calculus used in the argument. "Open-mindedness" refers to the disposition to form one's opinions by rational cognition and not on the basis of prejudices or emotions. A person is "discriminating" if she has the basic faculty of basic cognition and is able to organize respectively complex processes of cognition. (Cf. Lumer 1990, pp. 43 f.)

[vii] I would like to thank two anonymous referees for their valuable comments.

REFERENCES

Biro, J.I. (1987). A Sketch of an Epistemic Theory of Fallacies. In F.H. van Eemeren [et al.] (Eds.), *Argumentation, Analysis and Practics. Proceedings of the 1986 Amsterdam Conference on Argumentation* (pp. 65-73). Dordrecht: Foris.

Biro, J.I., & Siegel, H. (1992). Normativity, Argumentation and an Epistemic Theory of Fallacies. In F.H. van Eemeren [et al.] (Eds.), *Argumentation Illuminated* (pp. 85-103). Amsterdam: SicSat.

Carnap, R. (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Carnap, R. (1952). *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

Eells, E. (1982). *Rational decision and causality*. Cambridge: Cambridge U.P.

Feldman, R. (<1993> 1999). *Reason and Argument. 2nd Edition*. Upper Saddle River, N.J.: Prentice-Hall 1st ed. 1993; 2nd ed. 1999.

Gillies, D. (2000). Varieties of Propensity. *British Journal for the Philosophy of Science 51*, 807-835.

Goldman, A.I. (1999). *Knowledge in a Social World*. Oxford: Clarendon.

Goldman, A.I. (2003). An Epistemological Approach to Argumentation. *Informal Logic, 23*, 51-63.

Hacking, I. (1965). *The Logic of Statistical Inference*. Cambridge: Cambridge U.P.

Hacking, I. (2001). *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge U.P.

Hahn, U., & Oaksford, M. (2006). A Normative Theory of Argument Strength. *Informal Logic, 26*, 1-24.

Hahn, U., & Oaksford, M. (2007). The Rationality of Informal Argumentation: A Bayesian Approach to Reasoning Fallacies. *Psychological Review, 114*, 704-732.

Hájek, A. (<2002> 2009). Interpretations of Probability. *Stanford Encyclopedia of Philosophy*. Web publication: <http://plato.stanford.edu/entries/probability-interpret/>, first published 21 October 2002; substantial revisions 31 December 2009.

Hansson, S.O. (2004). Philosophical Perspectives on Risk. *Techne, 8*(1).

Humphrey, P. (1985). Why Propensities Cannot Be Probabilities. *Philosophical Review, 94*, 557-570.

Korb, K.B. (2004). Bayesian Informal Logic and Fallacy. *Informal Logic, 24*, 41-70.

Lumer, Ch. (1990). *Praktische Argumentationstheorie: Theoretische Grundlagen, praktische Begründung und Regeln wichtiger Argumentationsarten*. Braunschweig: Vieweg.

Lumer, Ch. (1991). Structure and Function of Argumentations: An Epistemological Approach to Determining Criteria for the Validity and Adequacy of Argumentations. In F.H. van Eemeren [et al.] (Eds.), *Proceedings of the Second International Conference on Argumentation* (pp. 98-107). Amsterdam: Sicsat.

Lumer, Ch. (2005a). The Epistemological Approach to Argumentation: A Map. *Informal Logic, 25*, 189-212.

Lumer, Ch. (2005b). The Epistemological Theory of Argument – How and Why? *Informal Logic, 25*, 213-243.

Mellor, D.H. (2005). *Probability: A Philosophical Introduction*. London / New York: Routledge.

Miller, D.W. (1994). *Critical Rationalism: A Restastement and Defence*. Chicago: Open Court.

Popper, K.R. (1959). The Propensity Interpretation of Probability. *British Journal for the Philosophy of Science, 10*, 25-42.

Ramsey, F.P. (<1926> 1931). Truth and Probability. (1926.) In F.P. Ramsey, *The Foundations of Mathematics and Other Logical Essays* (pp. 156-198), ed. by R. B.

Braithwaite. London: Routledge and Kegan Paul 1931.

Reichenbach, H. (<1935> 1949). *Wahrscheinlichkeitslehre: Eine Untersuchung über die logischen und mathematischen Grundlagen der Wahrscheinlichkeitsrechnung.* (1935.) – Enlarged English translation: *The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability.* Berkeley: University of California Press 1949.

Siegel, H., & Biro, J. (1997). Epistemic normativity, argumentation, and fallacies. *Argumentation, 11*, 277-292.

Venn, J. (<1866> 2006). *The Logic of Chance*. (1st ed. 1866.) Mineola, NY: Dover Publications 5th ed. 2006.

von Mises, R. (<1928> 1981). *Wahrscheinlichkeit, Statistik und Wahrheit*. Wien: J. Springer 1928. – 2nd revised English edition: *Probability, Statistics and Truth*. New York: Dover 1981.