

## **PROBABILITY AND CERTAINTY**

JONNY BLAMEY

KINGS COLLEGE LONDON

### ***Abstract***

Probability can be used to measure degree of belief in two ways: objectively and subjectively. The objective measure is a measure of the rational degree of belief in a proposition given a set of evidential propositions. The subjective measure is the measure of a particular subject's dispositions to decide between options. In both measures, certainty is a degree of belief 1. I will show, however, that there can be cases where one belief is stronger than another yet both beliefs are plausibly measurable as objectively and subjectively certain. In ordinary language, we can say that while both beliefs are certain, one belief is more certain than the other. I will then propose second, non probabilistic dimension of measurement, which tracks this variation in certainty in such cases where the probability is 1. A general principle of rationality is that one's subjective degree of belief should match the rational degree of belief given the evidence available. In this paper I hope to show that it is also a rational principle that the maximum stake size at which one should remain certain should match the rational weight of certainty given the evidence available. Neither objective nor subjective measures of certainty conform to the axioms of probability, but instead are measured in utility. This has the consequence that, although it is often rational to be certain to some degree, there is no such thing as absolute certainty.

### ***The No Certainty Principle and the Maximum Likelihood Principle***

Many Bayesians hold to the principle that a belief in a predictive empirical proposition should never be certain, or nearly certain, unless it is logically entailed

by the evidence (Good 1950, p. 36; Kaplan 1996, p. 93). Another principle is that you should proportion your belief to the evidence, and if the evidence consists in observed relative frequencies then you should believe a prediction to the degree that matches the relative frequencies, or to put it another way, to a degree that makes the observed frequencies retrospectively maximally likely (Fisher 1922, pp. 310 - 312; Popper 1974, p. 168; Good 1950, p.77).

These two principles, the ‘no certainty principle’ and the ‘maximum likelihood principle’, clash when all observations to date have been uniform, in other words, when the relative frequency has been 1 : 0. For example, if I have seen many ravens, all of which have been black, and my lover tells me that she has made a gift of a raven for me that is hidden in a casket, then the maximum likelihood principle tells me to be certain that the raven in the casket is black.

Predicting that the probability of being black given being a raven is 1 makes my previous observations most likely, so I should predict with probability 1 that the raven in the casket is black. However, the “no certainty principle” does not allow this. Since being black does not follow logically from being a raven, we should never assign probability 1 to black given raven. If we did, then it would be impossible, according to Bayesian conditionalization, to ever change our belief (Howson & Urbach 1989, pp. 284- 287; Williamson 2000, pp. 204, 218-220; Kaplan 1996, pp. 91 – 93). This clash of principles can be seen as a problem for statistics and mathematics, but it is clear that if we consider the philosophical problem of scepticism, we get this same duality of advice. It seems on one hand that we should be certain of many things, but on the other that we should be certain of nothing.

### ***Consumer and Producer Accounts of Belief***

These two principles are input driven, born out of a priori reasoning on the nature of evidential support. If we use Ruth Millikan’s (1993, p. 126) terminology, they are

both *producer* principles. Certainty is defined as a relation to a class of evidential support. If we are to resolve the dispute in a non circular fashion we need a consumer account. What is certainty *for*? In what way is certainty consumed by our mental economy? What does certainty *do*?

In decision theory, the output of degree of belief is usually defined in terms of the odds of an indifferent bet. In this sense, degrees of *confidence* are measured by probabilities. It is assumed that the more confident you are that a proposition is true, the greater the odds at which you would be prepared to bet. Certainty, on this view, is the highest degree of confidence. Certainty is when you would bet on  $p$  whatever the odds, and refrain from betting against  $p$  whatever the odds.

If odds are the only measure of subjective degree of belief, then once a subject is indifferent between a bet at 1 : 0 odds there is literally no way of being more certain, the measure not allow it, thereby giving certainty the character of infallibility. But as several people have pointed out in different contexts, once this threshold has been transcended, it is still possible to be more or less certain (Millikan 1993, p. 253; Williamson 2000, p. 99).

If a man publicly claims that his parrot is two hundred and fifty years old I would bet at any odds that he was speaking falsely and at no odds that he was telling the truth. In the consumer sense, in terms of output, it seems I am certain that his parrot is less than 250 years old. The maximum likelihood principle seems to endorse this, since assigning probability 1 to parrots dying before they reach 250 years old is the best way to account for the fact that I have never seen nor heard of any parrot approaching that age. But when a second man claims that his great grand father is still alive at the age of 350 years old, I find that I am even more certain of the negation of this second proposition. Yet how is it possible to

be even more certain than certain? Can there be something more probable than probability 1?

### *Ramsey's Consumer Measure*

Frank Ramsey in 1926 wrote a remarkable paper called 'Truth and Probability' where he presented a consumer theory of subjective probability. He compares degrees of belief to time intervals in physics. He argues that 'the degree of belief is just like a time interval; it has no precise meaning unless we specify more precisely how it is to be measured.' (Ramsey 1926, p. 63). He goes on to argue that in order to have a producer account of probability, we must have a precise measurement for what it is to have the corresponding degree of belief in those propositions, i.e. a corresponding consumer account.

After considering the problems associated with using the lowest odds at which a subject is prepared to bet on  $p$  as a measure of the subject's degree of belief that  $p$ , namely, the subject's reluctance or over eagerness to bet and the diminishing marginal utility of money, Ramsey developed a measure for utility based on preference. The subject's preferences between hypothetical worlds are ordered on an interval scale that I will indicate using capital letters A, B and C. Ramsey's genius was to transform the ordinal scale into an interval scale by using the concept of an ethically neutral proposition. An interval scale is one where, though zero and the unit is arbitrary, once the unit is decided, the interval between any two values is commensurable with the interval between any other two values. With an interval scale of utility, capitals 'A', 'B'...etc. can represent the utility to a subject of a set of worlds between which he is indifferent.

An interval scale can quantify the ratio of  $A - C$  to  $A - B$ . Time is also measured on an interval scale. For example, because time is measured on an interval scale, we can state clearly that the time difference between 2pm and 4pm is twice the

time difference between 5pm and 6pm. This is important in utility terms because it means we can give an exact value to the ratio of the utility difference between losing a bet and not taking the bet to the difference between winning and losing the bet.

### ***Ramsey's measure for degrees of belief***

Having established a method for measuring utility, Ramsey (1926, p.75) defined the following precise measure for a subject's degree of belief that p:

If the subject is indifferent between options:

1. A for certain, or
2. B if p and C if  $\sim p$ .

Then the subject's degree of belief that p is equal to:

$$\frac{A-C}{B-C}$$

He then demonstrated from this basic definition and constraints of consistency over preferences that subjective degrees of belief should obey the basic laws of probability.

### ***A Consumer model of Certainty***

On Ramsey's model, the consumer measure of certainty becomes the disposition to be indifferent between options 1 and 2 whenever  $A = B$ , since if one is certain that p the choice of options 1 and 2 becomes a straight choice between A for certain and B for certain. This is the prescription to act whenever there is any advantage at all to acting if p is true over not acting, regardless of the disadvantage of acting if p is false. C, which is the loss if the belief is false, has disappeared out of the equation. But C is still relevant to decision making.

Since we are looking for another dimension of measurement for degree of belief, we can take advantage of this hidden variable and use it to measure the level of

certainty when the degree of belief is at probability 1. Certainty is associated strongly with knowledge, and recent theorists have linked the truth of knowledge ascriptions to what is at stake in the situation. (Stanley 2005; DeRose 2002; Millikan, 1993 pp. 252 - 255). What is at stake in the situation can be defined precisely as the utility difference  $B - C$ , which is the difference in utility between acting if  $p$  and acting if  $\sim p$ . The more certain you are, the larger this value can be before electing not to act.

So we have got our output theory, or our consumer interpretation. I am more certain that the man is less than 350 years old than that the parrot is less than 250 in that the stakes on a bet would have to be a lot higher on the man before I would decline from accepting any odds. We can call the willingness to accept a bet at any odds a *guarantee*. Formally a guarantee can be defined as indifference between:

Option 1, A for certain, and

Option 2, B if  $p$  and C if not  $p$ , whenever  $A = B$ .

The amount of the guarantee will then simply be the difference between B and C, where C is less preferable than B. This amounts to C being a penalty for being wrong. At low stakes I would happily *guarantee* that the great grandfather is not 350yrs old (or is dead), in that I would be indifferent between option 1, not losing anything either way, and option 2, a penalty of £100, if the grandfather was 350 yrs old, and nothing, if he was younger than 350yrs old.

Since I am certain that the man is not 350yrs old, the two options have the same expected utility. I don't expect to have to pay £100 at all. The same goes for a £100 on the parrot being under 250yrs old. But as the penalty increases in value, I may opt out of the parrot guarantee sooner than I would opt out of the man guarantee. Doubts as to whether parrots could in fact live to incredible ages would begin to creep in at lower guarantees than doubts that there could be a 350yr old man. We

could in this fashion give an exact utility measure of my dispositional certainty in a proposition. It would amount to the precise threshold of loss at which I would cease to be indifferent between the two options. So at low stakes I am certain of both, but there is a range of stakes such that I would be uncertain that the parrot is not 250yrs old, but would remain certain that the man was not 350yrs old. The certainty threshold would be a utility measure and would not conform to the axioms of probability.

***Producer Theory for Probability: ratio of success to failure in experience***

What we need now is a producer interpretation, or an input theory. The most general and simple input theory for probability is that the degree of belief that an event type E will have property P is equal to the ratio of observed events type E with property P to observed events of type E with or without the property. If one in three men in your experience has been bald, the probability you should assign that the next man you meet is bald is 1/3. If all the men you have met have a heart, then the probability that the next man you meet will have a heart should be 1. This is a naïve frequency interpretation of probability. It is easy to show that these figures will conform to the axioms of probability.

There are many interesting questions as to how to type the predicted event, what to count as positive instances and so on, but the general principle remains the same. We can see that, as a general input theory for all epistemological issues, it is inadequate, for the very simple reason that it gives no extra weight to the sheer size of the sample from which the prediction is made. It also gives no extra weight to the quality of the sample, and the relevance of the sample to the event in question. For example, this input rule would prescribe the same actions whether i) I had only met six men in my life, two of whom were bald; ii) I knew that world wide one in three men are bald; or iii) I knew the next three men I was going to meet, but not in which order, and knew exactly one

of them to be bald. Intuitively these three samples are epistemically different, though each would recommend a degree of belief  $1/3$  that the next man I meet will be bald.

### ***Size Relevance and Depth of Experience Important***

So, when thinking about the degree of support an evidence sample lends to a prediction, it is not merely the ratio of cases that is relevant. There is also the size and relevance of the sample. In general, the broader and deeper one's experience, the better one's prediction. So a producer rule for certainty is that the greater the experience, the greater the certainty. Although the ratio of parrots under 250 years old to parrots in my experience, is equal to the ratio of men under 350 years old to men in my experience, (they are both 1 : 1); it is clear that my experience of men is greater in terms of relevance, depth and sheer number. If the input function for probabilistic degree of belief is the just a function of ratio, then there is room for non probabilistic variation in degree of certainty.

### ***Another Input Measure beyond Probability***

A fruitful place to look for such a measure is the cost of acquiring the evidence. Consider a repeatable experiment with two possible outcomes defined as successful or unsuccessful. The experience set can be expressed as the ratio of the number of successful trials to the total number of trials. When all the trials are successful we can predict with certainty (in the minimal sense) that a future trial will be successful. The degree of certainty can be defined simply as the number of trials. To convert this into a utility measure we can multiply the number of trials by the cost of performing a trial. In a case as simple as this, it is easy to see the logic. If the prediction of a trial demanded a high threshold, in that the penalty for being wrong was higher than the cost of performing more trials, then it would seem sensible to perform more trials. Whereas if the penalty for being wrong was less than the



cost of having performed the number of trials in experience, then it would seem sensible to go ahead and risk being certain. A similar line of thinking has been pursued by Good (1967):

“Our conclusion is that, in expectation, it pays to take into account further evidence, provided that the cost of collecting and using this evidence, although positive, can be ignored. In particular, we should use all the evidence *already* available, provided that the cost of doing so is negligible. With this proviso then, the principle of total evidence follows from the principle of rationality.”

In general Good is a proponent of the no certainty principle. His conclusion therefore does not apply to certainty. If we do not ignore the cost of “collecting and using the evidence” then collecting evidence will only pay if the advantage given by the information outweighs the cost. To make this decision we must have a value for the advantage given by the information. In an obscure note by Ramsey (1991, p. 285) entitled “Value, or weight of knowledge” Ramsey gives a formula for the value of knowledge that you already know, in other words the value of finding out that same information in relation to the expectation of advantage of an unknown proposition. If we take it that our current knowledge gives probability 1 to a prediction, we can use Ramsey’s formula to decide on the value of our current knowledge and whether it is adequate for the risk inherent in the prediction. The “value of knowledge” will of course be relative to interests. Given that there is no knowable difference between that which we know and that of which we are certain, (Ramsey 1991, p. 82) we now have both pieces of the jigsaw: an input measure for the value of certainty and an output measure in terms of the magnitude of the loss in error.

I know of a lot more ages of death of men than of parrots, and the cost of significantly increasing that knowledge is inestimably high with regard to men, but lower with

regard to parrots. (I could look into the life spans of parrots and increase my knowledge quite easily). So my certainty threshold *should* be as it is, higher with regard to the man than to the parrot. The literature on the interest relativity of knowledge ascriptions (Stanley 2005; DeRose 2002) would be much clarified by this model. That S appears to know that  $p$  at low stakes, but does not know that  $p$  at high stakes given equivalence of evidence would be explained by the certainty threshold lying between high stakes and low stakes. The high stakes situation would make seeking further evidence cost effective.

### ***Certainty Threshold Useful for Warrant Transmission***

The idea of a cut off stake size for certainty is especially pertinent when transmitting warrant through testimony. It is often the case that a more experienced and knowledgeable person will want to transmit the epistemic status of their prediction without having to say exactly what their experience consists in. For this reason there are many words in natural language which give an indication of whether the cut off stake size has been reached. Certainty is the most basic of these. It has often been pointed out that certainty is intuitively incompatible with less than 1 probability. The most famous illustration of this point is the lottery paradox whereby it feels unnatural and wrong to say that it is certain that a ticket will lose before the draw, however unlikely the ticket is to win (Kaplan 1996; Hawthorne *et al* 1999; Kyburg 1961). Yet certainty is on a sliding scale, there is nothing infelicitous in talking of increasing or decreasing levels of certainty. This paradox also applies to the term “sure”. Other words in this ball park are “knowledge” and “outright belief”. While not so clearly gradable, (Jason Stanley<sup>1</sup> denies that knowledge is gradable at all), both seem at once to denote a probability 1 and yet to be sensitive to non evidential features of the practical situation. Both Williamson and Millikan have noticed this link between probability 1 and variation due to stake size.

---

<sup>1</sup> This was from a talk he gave in St Andrews and London in 2006 called *Knowledge and Certainty*.

“Outright belief still comes in degrees, for one may be willing to use  $p$  as a premise in practical reasoning only when the stakes are sufficiently low” (Williamson 2000, p. 99)

“Unless the stakes are high, we do not require of the person who “knows that  $p$ ” because he has been told by so-and-so that he be ready to place extremely high stakes on the truth of  $p$ ” (Millikan 1993, p. 253)

The way to look at it is this: if you were in a betting shop and had to decide on whether to bet on  $p$ , there are two questions to be asked. “At what odds should you bet on  $p$ ?” and “how much should you bet?” These questions can be translated as: “how probable is  $p$ ?” and “how certain is  $p$ ?” The first figure will be a ratio that conforms to the axioms of probability calculus. The second will be an integer value with no absolute limit. Both these values have their place in ordinary talk, but in science people only use probability (of course other features of the data can be represented in some cases by statistical conventions like standard deviation, but in cases where the probability of an event is 1, standard deviation doesn’t have any meaning).

### ***Probabilism, second order Probabilities***

An alternative to the view I am putting forward here I shall for ease of expression call “Probabilism”. Probabilism is the view that any variation in prescriptive or descriptive degree of belief is a variation in probability that should conform to the axioms of probability calculus. A probabilist could object that the increase in the sample size simply increases the probability of the generalisation which issues the prediction. As the sample increases in size, Bayesian dynamics will ensure that the probability of the generalisation will tend towards 1 without ever reaching 1. This has the consequence that no predictions are certain since no generalisation is ever certain.

For example, suppose I have only ever seen and heard of black ravens. Each extra black raven increases my degree of belief in the hypothesis that all ravens are black. The more ravens I have seen, the greater my belief in this universal generalisation. Any singular prediction will have a probability conditional on the probability of the generalisation being true. This has the consequence that the probability of a prediction is always lower or equal to the probability that the generalisation is true. More formally, if the evidence “E” is counted as known such that E consists in a large number of sightings of black ravens and no sightings of other ravens; whereas the generalisation “H” is the hypothesis that all ravens are black, then  $(H|E) < 1$  by the no certainty rule.

If we know of a raven but we do not know its colour, then it seems that given E we should predict with a high probability that the unobserved raven is black (RB). What is  $P(RB | H)$ ? Surely it is the  $P(RB + H)/P(H)$  (Kaplan 1996 p.48). But in this case the  $P(RB | H)$  is less than or equal to  $P(H)$ . Of course RB might be true but H false. But in this case it is hard to see what relevance your degree of belief in the generalisation that all ravens are black has to your prediction that the raven will be black. The generalisation is redundant.

### ***Probabilism is Counterintuitive***

Probabilism is counterintuitive since a singular prediction is a lesser commitment than a generalisation. I am intuitively a lot more certain that the next person I meet has a heart than I am that all living humans have a heart. Secondly the generalisation is likely to be tailored to the prediction rather than the other way around. If I tell you that I am certain that that pill won't kill you I am not necessarily committing myself to a generalisation that all pills won't kill all people, or even that no pill of that kind ever kills anyone. Some aspirins may kill some people, some pills that find their way into aspirin bottles may kill anyone. What I am guaranteeing is that *this* aspirin won't kill *you*.

***Probabilism fails to distinguish between, probability 1 at low levels of certainty and probability of less than 1 at high levels of certainty***

I may be certain that if you go on this bungee jump you won't die on the grounds that a hundred people have been on this bungee jump and none of them have died. A probabilist says I shouldn't predict with probability 1 that you won't die on this bungee jump. I should instead leave some room for a possible increase in certainty. Let's say that I should predict with probability 0.999 that you will not die. But suppose there is another bungee jump where a thousand people have jumped, one of whom died? Should I predict that this bungee jump is *no more dangerous* than the one that was used less, but that has had a perfect safety record? What of a bungee jump where a hundred million people had jumped and one in every thousand had died? In this case I would be reasonably certain that the chance of dying was one in a thousand. In the first case the 0.999 safety figure is cooked up, plucked out of the air. Much better to say that the probability of dying is 0, *as far as we know*.

***Threshold Theories***

A plausible line of argument that a probabilist might give is that the term "certainty" as used by the folk is to be interpreted as a Bayesian degree of belief over some threshold less than 1 (Kaplan 1996, p. 93 – 101). But this is also problematic for the Bayesian who considers the degree to which evidence confirms a theory (T) to be the increase in probability that the theory is true conditional on the evidence. This has the consequence that the results of a scientific experiment have a probability 1. If the result of a scientific experiment was (e), and the prior probability of  $(T | e) = n$ , then the new probability of T should be n. This has the implicit assumption that if (e) is scientific evidence, then the probability  $(e) = 1$ . A reason to reject any threshold theory is that it would seem peculiar if the results of scientific experiments had a higher probability than certainty since even scientific observations are fallible. So if the results of scientific experiments have a probability 1, then propositions of which we are certain must have

a probability of at least 1. But there can be no probability higher than 1. Further more if we allow observations to have a probability of 1, and yet the threshold of certainty to be less than 1, we get an alarming failure of modus ponens.

Suppose for the sake of reductio that the threshold for certainty is 0.99. Suppose that  $P(h|e) = 0.99$  and is therefore certain and  $P(e) = 0.99$  and so certain. Then the probability (h) could be less than probability (e). For example if  $p(h) = p(h + e) = 0.9801$ . But this is absurd, since it has the consequence that (e) could be certain and the probability of (h|e) also certain and yet (h) be less than certain. For example let (h) be that Jones is mortal and (e) be that Jones is dead. Let us say that (e) is certain and (h|e) is certain and this means that they both have probability 0.99. Suppose also that it is certain that Jones is dead. It is possible, with the right prior probabilities, that the posterior probability that Jones is mortal should be 0.9801, and therefore not be certain. The absurd consequence is that it can be rational to be certain that Jones is dead, and to be certain that if Jones is dead then he is mortal, but to be not certain that he is mortal. These arguments are conclusive that if certainty implies any probability at all it implies probability 1. Probabilism therefore cannot explain varying degrees of certainty. This leaves variation in cut off stake size due to sample size and quality as the only account available.

### ***Generalisations are counterfactual supporting***

Yet another reason to reject probabilism is that the generalisations themselves must be counterfactual supporting since we can make predictions of counterfactual events of which we are certain. This has to be the case otherwise we could not use predictions in order to avoid doing actions with dangerous or fatal consequences. If I shoot myself in the head then I am certain that I will die, therefore I will not shoot myself in the head. Because generalisations have to be counterfactual supporting, the population over which they are quantified is infinite. Since any data set is going to be finite, the subjective probability that the generalisation is true can

never raise above 0 since the generalization can never be confirmed. (Popper 1974, p.63) Also, if the rationale behind the increase of probability of the generalisation due to increase in sample size is that the bigger the sample, the more likely it is to accurately represent the total population, then the rationale fails for infinite populations. (Strawson 1956, p. 253 -255) This is because the constitution of any sample of any size has a zero probability of matching the constitution of the total population since any number divided by infinity is zero.

### ***Probabilism as an Output Function is Insensitive to Risk***

So from the input side there is plenty of reason to endorse an independent measure of certainty which is a function of the size and quality of the sample, rather than to try and translate any increase in epistemic standing as an increase in a single probability figure. It is not hard to find criticisms of probabilism on the output side as well, and these can easily be tackled by having an independent certainty scale. The focus of the criticisms is that subjective probability is insensitive to the subject's attitude to risk and to the diminishing marginal utility of money. The most famous criticism is from the Allais problem (Allais 1953). Allais showed that intuitively non gambles are more valuable than gambles even when the expected utility is the same.

### ***Solution to Allais Problem through degrees of Certainty***

My strategy is to use  $B - C$ , as a separate measure of certainty.  $B$  is the value of an action conditional on  $p$  and  $C$  is the value of the same action conditional on  $\sim p$ . Think of  $B - C$  as representing what is at stake in assuming  $p$  with regard to a specific action. On my view there is no need to assume that the rational degree of belief will remain the same as  $B - C$  varies in magnitude. The assumption that the magnitude  $B - C$  has no affect on the rational degree of belief is demonstrably false, since if  $B - C$  is 0 then the truth or otherwise of  $p$  becomes irrelevant to the choice of whether to act or not, and the subject's degree of belief that  $p$  becomes

undefined by her choice of actions. This is a purely mathematical result since whenever the denominator is 0, the value of the fraction is undefined. The Allais problem exploits this and presents a series of options one of which has  $B - C$  as zero. It seems intuitively rational to favour the option where  $B - C$  is zero, especially when  $B$  and  $C$  are themselves of great positive utility. (Anecdotally Savage found himself preferring the sure gain over the higher expected utility, going against his own theory. However, instead of amending his theory, he amended his intuitions.) Seeing this mechanism it is easy to come up with Allais-type problems for subjective probability. Given the choice of either:

Option 1: WIN £1 million if two sixes are thrown on a fair die and WIN £1 million if two sixes are not thrown; or

Option 2: WIN £50 million pounds if two sixes are thrown on a fair die but LOSE £100 000 if two sixes are not thrown.

It would be a reckless person who chose option 2 even though the expected utility is much higher. Please ask yourself if you would really view it as more rational to go for a  $1/36$  chance of winning £50 million and a  $35/36$  chance of losing £100 000, over a certain million?

Also, if one could bet as much as one liked on the single toss of a coin at odds of 11:20 in one's favour, how much should one bet? Surely not every penny to one's name. Commentators are split on these issues. Ramsey (1926, p. 72), Friedman & Savage (1948) and Jeffrey (1970, pp. 161 - 2) were persuaded that these problems can be solved by recourse to a robust utility theory. Ramsey added the value  $A$  to counterbalance the subject's attitude to bets. He presumed that his utility theory would do the rest.

Others think these problems are decisive and that subjective probability theory should be seriously modified or abandoned (Kaplan 1996, p. 172; Allais 1953). These



latter tend to ignore the solution inherent in adding the counterbalancing value  $A$  for not accepting the bet. I am suggesting a middle ground. This is to drop the assumption that the rational bettor should accept the same range of odds whatever the magnitude of the difference between the loss and the gain. This assumption is counterintuitive, as the Allais problem shows. It is also demonstrably false since when  $B - C = 0$  then the bettor is rational to accept any odds.

Furthermore, it is not argued for by a Dutch book argument since there is no way a cunning bettor can have simultaneous bets at different stake sizes, since this would be in effect to have a single bet with the sum of the various stake sizes. (If I bet you £10 that it is raining and £100 that it is not raining, both at 1 : 1, I am not making two bets; I am making one bet of £90 that it is not raining). A final consideration in favour of dropping this assumption is that the truth of knowledge ascriptions seems to be both sensitive to practical interests and yet be negated by sub unitary probabilities. If probability itself varies with stake size, then the probability necessary for knowledge can be invariant at 1 while the knowledge ascriptions themselves vary with stake size.

### ***Conclusion***

The conflict between the maximum likelihood principle and the no certainty principle can be resolved if we consider the practical prescriptions given by certainty. One can rationally be certain when the stakes are low but become less than certain when the stakes get higher. The exact point at which one ceases to be certain is the *degree of certainty* as opposed to the degree of belief measured as a probability function. The degree of belief is the exact odds at which one would cease to bet. The degree of certainty is the exact stakes at which one would cease to bet at any odds.

The degree of certainty so defined is an output function, it is certainty seen from a consumer point of view. If the consumer needs two dimensions, then the producer

must produce two dimensions. Degrees of belief can be produced by considering the ratio of successes to failures in experience. Degrees of certainty can be produced by considering the size and relevance of the experience. The no certainty principle can then be seen to be referring to the maximum degree of certainty.

Given the nature of value and experience, there is no absolute certainty so defined, so the no certainty principle is correct. The no certainty principle is only incorrect in so far as it suggests that we should never assign a probability 1 to a singular prediction. With these two dimensions we can give a normative account of the interest relativity of knowledge ascriptions. We can also concede to the sceptic that we can never be absolutely certain of anything. But since it is only the kind of certainty that is relative to our interests that is necessary for knowledge, we do not have to concede to the sceptic that we don't know anything. We are and should be certain of many things which are of little or no importance, but of the things that really matter it is wise to keep an open mind.

### ***References***

- Allais, Maurice, 1953. "Le comportement de l'homme rationnel devant le risqué: critique des postulats et axiomes de l'école Américaine", *Econometrica* 21: 503 – 46
- DeRose, Keith, 2002. "Assertion, Knowledge, and Context", *Philosophical Review*, Vol. 111, No. 2. Apr. pp. 167-203
- Fisher, R.A., 1922. "On the Mathematical Foundations of Theoretical Statistics", *Philosophical Transaction of the Royal Society. A*, 222: 309 – 368
- Friedman, M & Savage, L. J., 1948. "The Utility Analysis of Choices Involving Risk" *Journal of Political Economy* 56, pp 279 -304
- Good, I.J., 1967. "On the Principle of Total Evidence", *British Journal for the Philosophy of Science*, 17

Hawthorne, James and Bovens, Luc., 1999. "The Preface, the Lottery and the Logic of Belief", *Mind*, Apr., Vol.108, pp. 430

Jeffrey, Richard C., 1970. "Dracula meets Wolfman: Acceptance VS. Partial belief", Swaine, M. ed., *Induction Acceptance and Rational Belief*, D. Reidel Publishing Company, Dordrecht Holland

Howson, C & Urbach, P, 1989. *Scientific Reasoning the Bayesian Approach*. Open Court Publishing Company

Kaplan, Mark, 1996. *Decision Theory as Philosophy*, MIT University Press

----- 2006. *Coming to terms with our human fallibility*, Epistemic Value Conference and Pre-Conference Workshop August 18th-20th, Stirling Management Centre

Kyburg, Henry, Jr. 1961. *Probability and the Logic of Rational Belief*. Wesleyan University Press

Millikan, Ruth, 1993. *White Queen Psychology*, MIT University Press

Ramsey, F. P., 1926. "Truth and Probability" *Philosophical Papers*, ed. D.H. Mellor. 1990. Cambridge University Press

----- F. P., 1991. *Notes of Philosophy, Probability and Mathematics* Ed. Galavotti, M. C. Bibliopolis - Edizioni di filosofia e scienze

Stanley, J., 2005. *Knowledge and Practical Interests*, Oxford University Press

Strawson, P. F., 1956. *Introduction to Logical Theory*, Methuen

Williamson, T., 2000. *Knowledge and its limits*, Oxford University Press