WEIGHT OF EVIDENCE

Sören Halldén

1. The certainty of a probability

In business, in politics, and indeed in all fields of practical activity, uncertainty is present, and there is an interest in the values which are dependent upon uncertainty. Suppose that a major investment is under consideration, and that the possibility of a war in the Middle East has to be included in the picture. Can one be rational in face of such a difficulty? A standard answer is provided by the principle of expected utility which permits calculation if one can say anything sensible about (1) the value of the undertaking in itself and (2) the probability of the conditions affecting it. If U is the value sought, U_1 the value of the action evaluated under the condition C, $U2$ the value under the condition not-C, and P the probability of C, then the following identity is asserted:

$U = P$ times $U1$, plus $(1-P)$ times $U2$.

This principle has a respected position in the discussion of economists and decision theorists. But it becomes questionable if one makes use of an observation of some philosophers. The latter is old; it became well-known through Keynes' Treatise on Probability, but it goes back to a succinct remark of Charles Sanders Peirce in a paper published in the nineteenth century:

"...to express the proper state of our belief, not one number but two are requisite, the first depending on the inferred probability, the second on the amount of knowledge on which that probability is based."{1}

The relevance of this point of view for the principle of expected utility was pointed out by Peter Gärdenfors in 1979.{2} The simple truth is that probability is not enough; it does not provide a full measure of certainty. It is one thing to have arrived at a probability estimate, and quite another to be willing to rely on it. If the foundations are dubious, this affects the solution.{3}

The problem area becomes even more complex if one accepts the doubts of a well-known skeptic, David Hume.{4} In the development of one's views a certain hesitation is often motivated, and perhaps there is nothing we can feel entirely and completely sure about. Original naive confidence may thus be devoured by higher level doubt. But what about the certainty we feel when a probability is being judged? Is this second level certainty to be devoured by a third level doubt?

The answer is, very fortunately, that belief-formation of the kind we are interested in here depends upon what one has decided. You are free to choose your premisses, and certainty is relative to the position chosen. If you choose to be naive, you can do so; you may be more particular if this is your wish. In the present context, I think it is wise to choose the middle road, and the scope of higher level doubts will be restricted. Doubts on the second level will be accepted, but not those on higher levels.{5} On the first level we have probability judgments which are comparatively naive, on the second epistemic judgments which concern the validity of first order judgments. A third level may be recognized, but this is a level where absolute confidence reigns - everything which belongs to the second level is accepted. The second level will therefore have a high degree of security. Let me call it ”Gärdenfors security”, or shorter ”G-security”. It is safe from the kind of doubts indicated by Peter Gärdenfors.

But what does it mean to affirm the ”reliability of a probability”? The certainty of a degree of uncertainty, what is that? Hesitatingly, Gärdenfors speaks about the ”probability of a probability”.{6] This idea plays an important part in his original paper and it is taken up later, still with some hesitation, in a series of publications by Gärdenfors, Robert Goldsmith, and Nils-Eric Sahlin.

The predicament is well-known. Probabilities are carefully scrutinized by scientists and by practical men; judgments concerning their trustworthiness are made; action is influenced by the judgments which are formed. The importance of such considerations is not to be doubted; the intuitions on which they are based seem secure; they cannot be pushed away. But what sense can one make of the formulations that suggest themselves? What is the certainty of a probability?

2. A pragmatic approach

If one should not speak about the probability of a probability here, a second order probability, what else can one do? Is there any other way out? A pragmatic perspective is close at hand, and this suggests a solution. Epistemic values are, of course, intimately tied to cognitive behaviour. They determine not only the premises chosen, but also the investigative steps which are classified as rewarding. They influence the expectations, hesitations and fears by which the investigator is guided.

A game is played in which the intrinsic interest of hypotheses and the applications which they permit are weighed against the difficulties of resolution. Rewards and punishments are involved, glory and humiliation, advancement and defeat, insight, confusion and error. When a decision is taken, the investigator is faced with two different alternatives: either he is right and gets the reward, or wrong and receives the punishment. This means that he makes a wager; epistemic values are mirrored by the bets which are accepted. At a G-secure level the relationship becomes uncomplicated. There we can say, quite simply, that the more certain one is, the more one is willing to risk. The stakes one is willing to offer are affected; the certainty in question may be measured in terms of the bets one is willing to accept.

The pragmatic approach is attractive, and it has its consequences. A betting order which corresponds to epistemic values is easily established, and in this way we are brought back to a solution in terms of probabilities. One does not have to be a subjectivist to see that there is a correlation between a betting order that is epistemically acceptable on one hand and probability on the other. A special case is that in which the investigator refuses to accept a bet. Then he does not regard the alternatives which are offered as definitely equal or, perhaps, as at all comparable; he confesses his inability to choose. A distinction between probabilistic and pragmatic equality is thus indicated. However, if we restrict ourselves to cases in which there is a difference in betting order, we have the right to speak about a difference in probabilistic order. Something similar might be said about situations in which betting equality is accepted. So if we restrict ourselves to such cases we are led back to the idea that the investigator is really dealing with the probability of a probability.

How sharp is it then, how precise? The pragmatic network is admittedly somewhat diffuse, and then the same should be said about the corresponding probabilistic network. But models which are precise may be constructed; precision is then recovered.

However, with Sahlin who has noted the possibility of this kind of opening, one may question the realism of this kind of approach.{7} Betting on a possibility of the form $P(H) = x$, you may be betting on something unverifiable. Verifying either H, or not-H, you have not verified the probability statement. Perhaps there is no pay-off. The question arises whether you can take the betting situation seriously. Whatever the risk, you do not have to pay, and there is nothing to earn.

I admit that probabilistic statements cannot be sharply falsified on the observational level. Yet they are sometimes proved, and sometimes disproved, and they are often tested, often confirmed and often disconfirmed. This should be sufficient for the person who is playing the game of theory development. Also, I cannot see that realism is ultimately decisive in this connection. Asked in the right way to formulate your bets in a situation that is nonrealistic you will react as if the situation were realistic. The experiments reported by Robert Goldsmith and Sahlin{8} have reinforced me in this view. Subjects were then concerned with ”imagined monetary sums”, but this did not seem to disturb them. The force of imagination should not be belittled.

It may therefore be hard to find an interpretation of second order certainty which is definitely nonprobabilistic.

3. The assumption of completeness

Secondary probabilities seem to me inavoidable. The question then arises why they should be left out of the discussion. Critical voices are heard, but secondary probabilities are not condemned as illusory or incorrect. Instead they are classified as primary probabilities in disguise. This idea was introduced by Max Woodbury, but presented to the public by Leonard Savage.{9} Some of the difficulties of the translation program of these two writers have been pointed out by Nils-Eric Sahlin,{10} but there is something one can add about the relationship of the program to certain common forms of reasoning. My belief is that the Woodbury-Savage model does not catch the multiple forms of ordinary thinking that concern ”weight of evidence”.

First, I want to say something about a kind of criticism of the translation program that is implicit in the approach of Gärdenfors and Sahlin. In order to fully understand the Woodbury-Savage idea one has to see how it is rooted in statistical thinking. The probability distributions to which the statistician is habituated are complete. Completeness is assumed, and quite correctly - yet we are here concerned with a form of idealization. Within pure theory the assumption is unobjectionable, but there is serious trouble if theory is applied within the province of ordinary rational argument. The probabilities one has to count with in the development of one’s views are often incomplete; the values accepted are not always precise; one is not concerned with an identity of the form $P(H)=x$, but a judgment of the form $P(H)>x$, or $P(H)>x$. One is then dealing with a minimum value. If the hypothesis H is under discussion, the scientific worker may be willing to grant that H has at least a certain probability. Or he may ascribe a minimum value to not-H. But he is much too sensible to expect that the minimum for H and the minimum for not-H add up to one. There is an interval between the two, and he is well aware of this. (On the pragmatic level, this is connected with a reluctance to accept certain bets.)

This has something to do with the secondary probabilities which might be involved in the way the investigator looks at things. If probabilities of this kind are introduced in the discussion, it is usually because one feels uncertain about the full identity; the secondary probability contains information concerning the reduction one is contemplating. Suppose that you are seriously ill, and that your doctor suggests surgery. He seems hesitating, and you ask him about the risk. - ”The mortality rate has been about one percent”, he has to say. ”But”, he adds, ”there have been some advances recently in the technique, so the odds might have improved considerably.” A secondary probability is hinted at here, and it leads to a reduction in the figure you have to take a stand to. Alternatively, the reduction may concern the opposite assumption, the assumption that you will have a chance of survival. - ”The local surgeon has no experience in this field of surgery, so perhaps we should wait”.

In both cases uncertainty is emphasized, and the uncertainty in question is not probabilistic, not measured by a number. It should be seen as a confession of sheer ignorance. The introduction of the secondary probability (which may well be incomplete) means that a primary probability which is complete, is reduced to an incomplete form. Or, one can say, it is replaced by ”a set of epistemically possible probability measures”, to use an expression which is central to the work of Gärdenfors and Sahlin.{11} A new element is added, a more basic kind of uncertainty introduced; there is a radical change in the position taken by the investigator.

4. Weight and the actual

The translation of the secondary to the primary envisaged by Woodbury is acceptable only if one takes completeness for granted. But there are also other reasons for the failure of the program. One of the preconditions of translation is that the probabilities involved are defined in the same way, that they have the same meaning on both levels. Sahlin has pointed to the importance of this condition,{12} but the general theme may be developed further.

It is well-known that ”probability is always relative to evidence”. But when a probability is used in an argument, a specific foundation is not always indicated. The content is silently taken for granted. In the Woodbury case a tacit assumption is that the two probabilities, the secondary and the primary, refer to the same basis. But it is questionable if this is always realistic.

The purpose of the formal treatment is to shed light on the structure of ordinary reasoning. But suppose that you first formulate a primary probability judgment, and that afterwards a secondary judgment is added that concerns the former. Will it have the same basis? Isn’t it quite likely that the latter judgment will be based upon some new facts which you have just remembered, or discovered, facts which cast a new light on the problem you are confronted with? Perhaps you are now paying attention to something you knew already at the beginning. By a rational reconstruction of the argument, this should be treated as an addition to the original basis. The probability calculated builds on something previously disregarded. It is also possible that you have discovered something which is completely novel.

The fundamental issue is then if the primary probability can be regarded as stable after the scrutiny of the new evidence. A change to something numerically different, or a confession of ignorance of the type considered in sec 3 might be contemplated. Will it survive the confrontation with the facts? This is a question which concerns the stability after a moment of reconsideration. I will call it the ”stability ex post”. (”Stability ex ante” will come later, in sec 5.)

Let us turn back to the medical example cited above. You are first informed about the statistics concerning previous efforts in the field. So many have survived a specific form of surgical treatment, so many have died, and this provides the primary probability. The figure is based on what one knows concerning previous operations. Later, some extra information is added, concerning new technique, or concerning the competence of the local surgeon. The secondary judgment provides a correction of the primary one, but the foundations have changed. You know more, and this affects the meaning of the statement. A Woodbury reduction will then no longer make any sense at all.

Is this too easy a way out? The claim I make is that the secondary level statement, in the ordinary case, is based on additional information. And it is quite clear that this idea is tied to the presentation of the original problem, the problem of the weight of an argument. The formulation used by

Keynes when he explains the meaning of ”the weight of an argument” should be quoted here: ”As the relevant evidence at our disposal increases, the magnitude of the probability may either decrease or increase, according as the new knowledge strengthens the unfavourable or the favourable evidence; but something seems to have increased in either case, - we have a more substantial basis upon which to rest our conclusion. I express this by saying that an accession of new evidence increases the weight of an argument.”{13}

This means that the secondary judgment should be seen as a correction of the primary judgment. This is replaced by the secondary judgment. It can do this in virtue of the principle that if the additional information I is available, and P(H/G.I) is accepted, then P(H/G) should be disregarded. The principle in question is pragmatic in nature; the rule behind it is that one has to make use of what is at hand. It is certainly more reasonable to build on what one knows than on what one does not know.

5. Weight and possibility

It has been emphasized above that the amalgamation of two probabilities into a single number requires a special kind of relationship between the two. Attention was given to the case in which one tries to amalgamate P(H/G.I) and P(H/G), and it was easily seen that a pair of this nature is not suitable material for an amalgamation of the kind intended.

Let us now consider a different pair of probabilities where amalgamation becomes equally impossible. A new type of change is introduced which concerns the secondary probability. In the last section attention was paid to the basis, but let us now focus on the theme. Suppose that you are forming a judgment on a primary probability, expressed by the proposition P(H) = x. This constitutes a point of departure, and it will, in the following, be abbreviated D. In the Woodbury-Savage context it is presumed that the content of the secondary evaluation is just D - the judgment is of the form P(D) = y. But if one turns to ordinary argumentative contexts in which primary probabilities are judged one can see that a judgment of this kind has a higher degree of complexity, at least sometimes. A characterization of D, F(D), dealing with the robustness or stability of D, is involved.

The question of the validity of a probabilistic relationship D may be actualized at various stages in an investigation. At the end it may conform to the model treated in the Woodbury-Savage text (with a reservation to be formulated below). The theme of the secondary probability asked for is then just D. But at the beginning the approach will be different. Prognosis will be of vital interest; the question will be whether D will stand up to the tests that are to follow. The numerical content should perhaps be modified; a reduction in scope is perhaps requisite.

Inquiry often follows a standard pattern. There is a conventional procedure; observations are collected; measurements are made; experiments are carried through, experts are consulted. The inevitable question will then be: What are the chances that present beliefs will survive the evidential ordeal? Think here of the district attorney who has ordered the arrest of a suspicious character, a local hoodlum, but is well aware of the fact that twenty witnesses are waiting to be interrogated, and that one of them may provide the arrested man with an unbreakable alibi. The probability interesting him will concern a proposition of the form F(D), a proposition asserting the survival capacity of the preliminary guess D. P(F(D)) is not without connection with P(D), but the two have to be distinguished. Something which could be called ”robustness” or ”stability”, is in focus, when the question of P(F(D)) is raised.{14} The latter value could be called the ”stability ex ante”.

But let us return to the situation at the end of inquiry. Can we be sure that it conforms to the Woodbury-Savage model? At the end of inquiry it is both reasonable and appropriate to ask oneself whether one has really come to the end. Perhaps there are some loose ends which have to be tied up? What happens if one decides to go on? And this is a question which concerns the stability ex ante of the primary probability. The secondary probability is also here of the form P(F(K)). Once again the question concerns survival.

Is this a respectable interpretation of the problem of the ”weight of an argument”, as this is treated in the literature? Let me go back to an authoritative text, Keynes' Treatise on Probability. The author takes up the question ”to what point the strengthening of an argument's weight by increasing the evidence ought to be pushed.” He goes on: ”We may argue that, when our knowledge is slight but capable of increase, the course of action, which will, relative to such knowledge, probably produce the greatest amount of good, will often consist in the acquisition of more knowledge. But there clearly comes a point when it is no longer worth while to spend trouble, before acting, in the acquisition of further information, and there is no evident principle by which to determine how far we ought to carry our maxim of strenghtening the weight of our argument."{15}

The question of the stability ex ante is very much present in this discussion of the topic.

6. Multidimensionality

Let us now go back to the quotation from which I started: "...to express the proper state of our belief, not one number, but two are requisite..." Something which could be called "multidimensionality" is suggested here. The state of belief consists of a single judgment which can be characterized within two dimensions. It is like saying about Tom that he has blue eyes and is tall. I think that the idea of multidimensionality might be retained, but one has to emphasize that each of the dimensions is dependent upon the correctness of other judgments, and one has to add that we have no need to restrict ourselves to a two-dimensional account.

The point is that "weight" might be split up in at least two dimensions. There is a naive original judgment D, ascribing probability to a proposition. Various comments might be made which concern D. One provides a comment on D which is based on some extra information; the content concerns stability ex post. Another has something to say about the stability ex ante of D. These comments are to a certain degree independent of each other, and, if you want to, you can say that each one of them creates a new dimension. And they may both be of some importance.

Notes

I am indebted to Nils-Eric Sahlin for valuable comments.

1. Peirce, Collected Papers, vol 2, p 421.
2. Gärdenfors 1979.
3 Moreover, the decision maker has to take into consideration how certain he is, not only of the probability P, but also of the values U1 and U2.
4. Treatise, book 1, part 4, sec 1.
5. Sahlin 1983, p 96, emphasizes that "very little would be gained by considering higher than second order probabilities."
6. Gärdenfors 1979, p 172.
7. Sahlin 1983, p 97.
8. Goldsmith and Sahlin 1982. See especially p 462, 465.
9. Savage 1954, p 58.
10. Sahlin 1983, p 96; Sahlin 1993, p 25-26.

11. Gärdenfors and Sahlin 1982, p 317, 319. The possibility of complete ignorance with respect to the epistemic value of the measures in question is also mentioned by these writers.
12. Sahlin 1993, p 95-96.
13. Keynes 1921, Ch 6, p 71.
14. Cf Hansson 1983, p 89-97, and Halldén 1983.
15. Keynes 1921, Ch 6, p 77. Cf Levi 1967, Ch 9.

Bibliography

Gärdenfors, Peter. 1979. Forecasts, decisions and uncertain probabilities. Erkenntnis, vol 14, p 159-81.
 - and Sahlin, Nils-Eric. Unreliable probabilities, risk taking and decision making. P Gärdenfors and N-E Sahlin (eds). Decision, Probability and Utility, p 313-34. Cambridge University Press 1988. Originally published in Synthese, vol 53 (1982), p 361-86.
 - 1983. Decision making with unreliable probabilities. British Journal of Mathematical and Statistical Psychology, vol 36, p 240-51.
Goldsmith, Robert W and Sahlin, Nils-Eric. 1982. The role of second-order probabilities in decision making. P Humphreys, O Svenson and A Vàri (eds), Analysing and Aiding Decision Processes. Budapest.
Halldén, Sören. 1983. Probability logic is a dangerous tool. Second Scandinavian Congress on Forensic Science, p 55-62. Linköping University. Linköping.
Hansson, Bengt. 1983. Epistemology and evidence. In P Gärdenfors, B Hansson and N-E Sahlin (eds), Evidentiary Value. Library of Theoria, vol 15. Thales, Stockholm.
Peirce, Ch S. 1932. Collected Papers, vol 2. Hartshorne and Weiss (eds). Belknap. Cambridge, Mass.
Keynes, John Maynard. 1921. A Treatise on Probability. Macmillan, London 1963.
Levi, Isaac. 1967. Gambling with Truth. MIT, Cambridge, Mass.
Sahlin, Nils-Eric. 1983. On second order probabilities and the notion of epistemic risk. B P Stigum and F Wenstrup (eds), Foundations of Utility and Risk Theory with Applications. Reidel. Dordrecht.
 - 1993. On higher order beliefs. J-P Dubucs (ed), Philosophy of Probability, p 13-34. Kluwer. Dordrecht.
Savage, Leonard J. 1954. The Foundations of Statistics. Wiley, New York.