



Rational Decisions

Author(s): I. J. Good

Source: *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 14, No. 1 (1952), pp. 107-114

Published by: Blackwell Publishing for the Royal Statistical Society

Stable URL: <http://www.jstor.org/stable/2984087>

Accessed: 26/11/2009 12:01

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Royal Statistical Society and Blackwell Publishing are collaborating with JSTOR to digitize, preserve and extend access to *Journal of the Royal Statistical Society. Series B (Methodological)*.

<http://www.jstor.org>

RATIONAL DECISIONS*

By I. J. GOOD

[Received October, 1951]

SUMMARY

THIS paper deals first with the relationship between the theory of probability and the theory of rational behaviour. A method is then suggested for encouraging people to make accurate probability estimates, a connection with the theory of information being mentioned. Finally Wald's theory of statistical decision functions is summarised and generalised and its relation to the theory of rational behaviour is discussed.

1. *Introduction*

I am going to discuss the following problem. Given various circumstances, to decide what to do. What universal rule or rules can be laid down for making rational decisions? My main contention is that our methods of making rational decisions should not depend on whether we are statisticians. This contention is a consequence of a belief that consistency is important. A few people think there is a danger that over-emphasis of consistency may retard the progress of science. Personally I do not think this danger is serious. The resolution of inconsistencies will always be an essential method in science and in cross-examinations. There may be occasions when it is best to behave irrationally, but whether there are should be decided rationally.

It is worth looking for unity in the methods of statistics, science and rational thought and behaviour; first in order to encourage a scientific approach to non-scientific matters, second to suggest new statistical ideas by analogy with ordinary ideas, and third because the unity is aesthetically pleasing.

Clearly I am sticking my neck out in discussing this subject. In most subjects people usually try to understand what other people mean, but in philosophy and near-philosophy they do not usually try so hard.

2. *Scientific Theories*

In my opinion no scientific theory is really satisfactory until it has the following form:

(i) There should be a very precise set of axioms from which a purely abstract theory can be rigorously deduced. In this abstract theory some of the words or symbols may remain undefined. For example, in projective geometry it is not necessary to know what points, lines and planes are in order to check the correctness of the theorems in terms of the axioms.

(ii) There should be precise rules of application of the abstract theory which give meaning to the undefined words and symbols.

(iii) There should be suggestions for using the theory, these suggestions belonging to the technique rather than to the theory. The suggestions will not usually be as precisely formulated as the axioms and rules.

The adequacy of the abstract theory cannot be judged until the rules of application have been formulated. These rules contain indications of what the undefined words and symbols of the abstract theory are all about, but the indications will not be complete. It is the theory as a whole (i.e., the axioms and rules combined) which gives meaning to the undefined words and symbols. It is mainly for this reason that a beginner finds difficulty in understanding a scientific theory.

It follows from this account that a scientific theory represents a decision and a recommendation to use language and symbolism in a particular way (and possibly also to think and act in a

* This paper is based on a lecture delivered to the Royal Statistical Society on September 22nd, 1951, as a contribution to the week-end conference at Cambridge.

particular way). Consider, for example, the principle of conservation of energy (or energy and matter). Apparent exceptions to the principle have been patched up by extending the idea of energy, to potential energy, for example. Nevertheless the principle is not entirely tautological.

Some theoreticians formulate theories without specifying the rules of application, so that the theories cannot be understood at all without a lot of experience. Such formulations are philosophically unsatisfactory.

Ordinary elementary logic can be regarded as a scientific theory. The recommendations of elementary logic are so widely accepted and familiar, and have had so much influence on the educated use of language, that logic is often regarded as self-evident and independent of experience. In the empirical sciences the selection of the theories depends much more on experience. The theory of probability occupies an intermediate position between logic and empirical sciences. Some people regard any typical theory of probability as self-evident, and others say it depends on experience. The fact is that, as in many philosophical disputes, it is a question of degree; the theory of probability does depend on experience, but does not require much more experience than does ordinary logic. There are a number of different methods of making the theory seem nearly tautological by more or less *a priori* arguments. The two main methods are those of "equally probable cases" and of limiting frequencies. Both methods depend on idealizations, but it would be extremely surprising if either method could be proved to lead to inconsistencies. When actually estimating probabilities, most of us use both methods. It may be possible in principle to trace back all probability estimates to individual experiences of frequencies, but this has not yet been done. Two examples in which beliefs do not depend in an obvious way on frequencies are (i) the estimation of the probability that a particular card will be drawn from a well-shuffled pack of 117 cards; (ii) the belief which newly-born piglings appear to have that it is a good thing to walk round the mother-pig's leg in order to arrive at the nipples. (This example is given for the benefit of those who interpret a belief as a tendency to act.)

3. *Degrees of Belief*

I shall now make twelve remarks about degrees of belief.

(i) I *define* the theory of probability as the logic of degrees of belief. Therefore degrees of belief, either subjective or objective, must be introduced. Degrees of belief are assumed (following Keynes) to be partially ordered only, i.e., some pairs of beliefs may not be comparable.

(ii) F. Y. Edgeworth, Bertrand Russell and others use the word "credibilities" to mean objective rational degrees of belief. A credibility has a definite but possibly unknown value. It may be regarded as existing independently of human beings.

(iii) A subjective theory of probability can be developed without assuming that there is necessarily a credibility of E given F for every E and F (where E and F are propositions). This subjective theory can be applied whether credibilities exist or not. It is therefore more general and economical not to assume the existence of credibilities as an axiom.

(iv) Suppose Jeffreys is right that there is a credibility of E given F , for every E and F . Then either the theory will tell us what this credibility is, and we must adjust our degree of belief to be equal to the credibility. Or on the other hand the theory will not tell us what the credibility is, and then not much is gained, except perhaps a healthier frame of mind, by supposing that the credibility exists.

(v) A statistical hypothesis H is an idealized proposition such that for some E , $P(E | H)$ is a credibility with a specified value. Such credibilities may be called "tautological probabilities".

(vi) There is an argument for postulating the existence of credibilities other than tautological probabilities, namely that probability judgments by different people have some tendency to agree.

(vii) The only way to assess the cogency of this argument, if it can be assessed at all, is by the methods of experimental science whose justification is by means of a subjective theory.

(viii) My own view is that it is often quite convenient to accept the postulate that credibilities exist, but this should be regarded as a suggestion rather than an axiom of probability theory.

(ix) This postulate is useful in that it enables other people to do some of our thinking for us. We pay more attention to some people's judgment than to others'.

(x) If a man holds unique beliefs it is possible that everybody else is wrong. If we want him to abandon some of his beliefs we may use promises, threats, hypnotism and suggestion, or we may prefer the following more rational method: By asking questions we may obtain information about his beliefs. Some of the questions may be very complicated ones, of the form, "I put it to you that the following set of opinions is cogent: . . .". We may then show, by applying a subjective theory of probability, that the beliefs to which the man has paid lip-service are not self-consistent.

(xi) Some of you may be thinking of the slogan "science deals only with what is objective". If the slogan were true there would be no point for scientific purposes in introducing subjective judgments. But actually the slogan is false. For example, intuition (which is subjective) is the main instrument of original scientific research, according to Einstein. The obsession with objectivity arises largely from the desire to be convincing in published work. There are, however, several activities in which it is less important to convince other people than to find out the truth for oneself. There is another reason for wanting an objective theory, namely that there is a tendency to wishful thinking in subjective judgments. But objectivity is precisely what a subjective theory of probability is for: its function is to introduce extra rationality (and therefore objectivity) into your degrees of belief.

(xii) Once we have decided to objectify a rational degree of belief into a credibility it begins to make sense to talk about a degree of belief concerning the numerical value of a credibility. It is possible to use probability type-chains (to coin a phrase) with more than two links, such as a degree of belief equal to $\frac{1}{2}$ that the credibility of H is $\frac{1}{3}$ where H is a statistical hypothesis such that $P(E|H) = \frac{1}{4}$. It is tempting to talk about reasonable degrees of belief of higher and higher types, but it is convenient to think of all these degrees of belief as being of the same kind (usually as belonging to the same body of beliefs in the sense of section 4) by introducing *propositions* of different kinds. In the above example the proposition which asserts that the credibility of H is $\frac{1}{3}$ may itself be regarded as a statistical hypothesis "of type 2". Our type-chains can always be brought back ultimately to a subjective degree of belief. All links but the first will usually be credibilities, tautological or otherwise.

4. *Utilities*

The question whether utilities should be regarded as belonging to the theory of probability is very largely linguistic. It therefore seems appropriate to begin with a few rough definitions.

Theory of reasoning: A theory of logic plus a theory of probability.

Body of beliefs: A set of comparisons between degrees of belief of the form that one belief is held more firmly than another one, or if you like a set of judgments that one probability is greater than (or equal to) another one.

Reasonable body of beliefs: A body of beliefs which does not give rise to a contradiction when combined with a theory of reasoning.

A reasonable degree of belief is one which occurs in a reasonable body of beliefs. A *probability* is an expression of the form $P(E|F)$ where E and F are propositions. It is either a reasonable degree of belief "in E given F ", or else it is something introduced for formal convenience. Degrees of belief may be called "probability estimates".

Principle of rational behaviour: The recommendation always to behave so as to maximize the expected utility per time unit.

Theory of rational behaviour: Theory of reasoning plus the principle of rational behaviour.

Body of decisions: A set of judgments that one decision is better than another. Hypothetical circumstances may be considered as well as real ones (just as for a body of beliefs).

Reasonable body of decisions: A body of decisions which does not give rise to a contradiction when combined with a theory of rational behaviour.

A reasonable decision is one which occurs in a reasonable body of decisions.

We see that a theory of reasoning is a recommendation to think in a particular way while a theory of rational behaviour is a recommendation to act in a particular way.

Utility judgments may also be called "value judgments". The notion of utility is not restricted to financial matters, and even in financial matters utility is not strictly proportional to financial gain. Utilities are supposed to include all human values such as, for example, scientific interest. Part of the definition of utility is provided by the theory of rational action itself.

It was shown by F. P. Ramsey* how one could build up the theory of probability by starting from the principle of maximizing expected utilities. L. J. Savage has recently adopted a similar approach in much more detail in some unpublished notes. The main argument for developing the subject in the Ramsey-Savage manner is that degrees of belief are only in the mind or expressed verbally, and are therefore not immediately significant operationally in the way that behaviour is. Actions speak louder than words. I shall answer this argument in four steps:

(i) It is convenient to classify knowledge into subjects which are given names and are discussed without very much reference to the rest of knowledge. It is possible, and quite usual, to discuss probability with little reference to utilities. If utilities are introduced from the start, the axioms are more complicated and it is debatable whether they are more "convincing". The plan which appeals to me is to develop the theory of probability without much reference to utilities, and then to adjoin the principle of rational behaviour in order to obtain a theory of rational behaviour. The above list of definitions indicates how easily the transition can be made from a theory of probability to a theory of rational behaviour.

(ii) People's value judgments are, I think, liable to disagree more than their probability judgments. Values can be judged with a fair amount of agreement when the commodity is money, but not when deciding between, say, universal education and universal rowing, or between your own life and the life of some other person.

(iii) The principle of maximizing the expected utility can be made to look fairly reasonable in terms of the law of large numbers, provided that none of the utilities are very large. It is therefore convenient to postpone the introduction of the principle until after the law of large numbers has been proved.

(iv) It is not quite clear that infinite utilities cannot occur in questions of salvation and damnation (as suggested, I think, by Pascal), and expressions like $\infty - \infty$ would then occur when deciding between two alternative religions. To have to argue about such matters as a necessary preliminary to laying down any of the axioms of probability would weaken the foundations of that subject.

5. *Axioms and Rules*

The theory of probability which I accept and recommend is based on six axioms, of which typical ones are—

A1. $P(E | F)$ is a non-negative number (E and F being propositions).

A4. If E is logically equivalent to F then $P(E | G) = P(F | G)$, $P(G | E) = P(G | F)$.

There is also the possible modification—

A4'. If you have proved that E is logically equivalent to F then $P(E | G) = P(F | G)$, etc. (The adoption of A4' amounts to a weakening of the emphasis on consistency and enables you to talk about the probability of purely mathematical propositions.)

The main rule of application is as follows: Let $P'(E | F) > P'(G | H)$ mean that you judge that your degree of belief in E given F (i.e., if F were assumed) would exceed that of G given H . Then in the abstract theory you may write $P(E | F) > P(G | H)$ (and conversely).

Axiom A1 may appear to contradict the assumption of section 3 that degrees of belief are only partially ordered. But when the axioms are combined with the above rule of application it becomes clear that we cannot necessarily effect the comparison between any pair of beliefs. The axioms are therefore stronger than they need be for the applications. Unfortunately if they are weakened they become much more complicated.†

6. *Examples of Suggestions*

(i) Numerical probabilities can be introduced by imagining perfect packs of cards perfectly shuffled, or infinite sequences of trials under essentially similar conditions. Both methods are idealizations, and there is very little to choose between them. It is a matter of taste: that is why there is so much argument about it.

* *The Foundations of Mathematics* (London, 1931).

† See B. O. Koopman, "The axioms and algebra of intuitive probability", *Annals of Math.*, **41** (1940), 269–92.

(ii) Any theorem of probability theory and anybody's methods of statistical inference may be used in order to help you to make probability judgments.

(iii) If a body of beliefs is found to be unreasonable after applying the abstract theory, then a good method of patching it up is by being honest (using unemotional judgment). (This suggestion is more difficult to apply to utility judgments because it is more difficult to be unemotional about them.)

7. *Rational Behaviour*

I think that once the theory of probability is taken for granted, the principle of maximizing the expected utility per unit time is the only fundamental principle of rational behaviour. It teaches us, for example, that the older we become the more important it is to use what we already know rather than to learn more.

In the applications of the principle of rational behaviour some complications arise, such as—

(i) We must weigh up the expected time for doing the mathematical and statistical calculations against the expected utility of these calculations. Apparently less good methods may therefore sometimes be preferred. For example, in an emergency, a quick random decision is better than no decision. But of course theorizing has a value apart from any particular application.

(ii) We must allow for the necessity of convincing other people in some circumstances. So if other people use theoretically inferior methods we may be encouraged to follow suit. It was for this reason that Newton translated his calculus arguments into a geometrical form in the *Principia*. Fashions in modern statistics occur partly for the same reason.

(iii) We may seem to defy the principle of rational action when we insure articles of fairly small value against postal loss. It is possible to justify such insurances on the grounds that we are buying peace of mind, knowing that we are liable to lapse into an irrational state of worry.

(iv) Similarly we may take on bets of negative expected financial utility because the act of gambling has a utility of its own.

(v) Because of a lack of precision in our judgment of probabilities, utilities, expected utilities and "weights of evidence" we may often find that there is nothing to choose between alternative courses of action, i.e., we may not be able to say which of them has the larger expected utility. Both courses of action may be reasonable and a decision may then be arrived at by the operation known as "making up one's mind". Decisions reached in this way are not usually reversed, owing to the negative utility of vacillation. People who attach too large a value to the negative utility of vacillation are known as "obstinate".

(vi) Public and private utilities do not always coincide. This leads to ethical problems.

Example.—An invention is submitted to a scientific adviser of a firm. The adviser makes the following judgments:

- (1) The probability that the invention will work is p .
- (2) The value to the firm if the invention is adopted and works is V .
- (3) The loss to the firm if the invention is adopted and fails to work is L .
- (4) The value to the adviser personally if he advises the adoption of the invention and it works is v .
- (5) The loss to the adviser if he advises the adoption of the invention and it fails to work is l .
- (6) The losses to the firm and to the adviser if he recommends the rejection of the invention are both negligible, because neither the firm nor the adviser have rivals.

Then the firm's expected gain if the invention is accepted is $pV - (1 - p)L$ and the adviser's expected gain in the same circumstances is $p v - (1 - p)l$. The firm has positive expected gain if $p/(1 - p) > L/V$, and the adviser has positive expected gain if $p/(1 - p) > l/v$. If $l/v > p/(1 - p) > L/V$, the adviser will be faced with an ethical problem, i.e., he will be tempted to act against the interests of the firm. Of course real life is more complicated than this, but the difficulty obviously arises. In an ideal society public and private expected utilities would always be of the same sign.

What can the firm do in order to prevent this sort of temptation from arising? In my opinion the firm should ask the adviser for his estimates of p , V and L , and should take the onus of the

actual decision on its own shoulders. In other words, leaders of industry should become more probability-conscious.

If leaders of industry did become probability-conscious there would be quite a reaction on statisticians. For they would have to specify probabilities of hypotheses instead of merely giving advice. At present a statistician of the Neyman-Pearson school is not permitted to talk about the probability of a statistical hypothesis.

8. Fair Fees

The above example raises the question of how a firm can encourage its experts to give fair estimates of probabilities. In general this is a complicated problem, and I shall consider only a simple case and offer only a tentative solution. Suppose that the expert is asked to estimate the probability of an event E in circumstances where it will fairly soon be known whether E is true or false (e.g., in weather forecasts).

It is convenient at first to imagine that there are two experts A and B whose estimates of the probability of E are $p_1 = P_1(E)$ and $p_2 = P_2(E)$. The suffixes refer to the two bodies of belief, and the "given" propositions are taken for granted and omitted from the notation. We imagine also that there are objective probabilities, or credibilities, denoted by P . We introduce hypotheses H_1 and H_2 where H_1 (or H_2) is the hypothesis that A (or B) has objective judgment. Then

$$p_1 = P(E | H_1), p_2 = P(E | H_2).$$

Therefore, taking " H_1 or H_2 " for granted, the factor in favour of H_1 (i.e., the ratio of its final to its initial odds) if E happens is p_1/p_2 . Such factors are multiplicative if a series of independent experiments are performed. By taking logs we obtain an additive measure of the difference in the merits of A and B , namely $\log p_1 - \log p_2$ if E occurs or $\log(1 - p_1) - \log(1 - p_2)$ if E does not occur. By itself $\log p_1$ (or $\log(1 - p_1)$) is a measure of the merit of a probability estimate, when it is theoretically possible to make a correct prediction with certainty. It is never positive, and represents the amount of information lost through not knowing with certainty what will happen.

A reasonable fee to pay to an expert who has estimated a probability as p_1 is $k \log(2p_1)$ if the event occurs and $k \log(2 - 2p_1)$ if the event does not occur. If $p_1 > \frac{1}{2}$ the latter payment is really a fine. (k is independent of p_1 but may depend on the utilities. It is assumed to be positive.) This fee can easily be seen to have the desirable property that its expectation is maximized if $p_1 = p$, the true probability, so that it is in the expert's own interest to give an objective estimate. It is also in his interest to collect as much evidence as possible. Note that no fee is paid if $p_1 = \frac{1}{2}$. The justification of this is that if a larger fee were paid the expert would have a positive expected gain by saying that $p_1 = \frac{1}{2}$, without looking at the evidence at all. If the class of problems put to the expert have the property that the average value of p is x , then the factor 2 in the above formula for the fee should be replaced by $x^{-x}(1-x)^{-(1-x)} = b$, say.* Another modification of the formula should be made in order to allow for the diminishing utility of money (as a function of the amount, rather than as a function of time). In fact if Daniel Bernouilli's logarithmic formula for the utility of money is assumed, the expression for the fee ceases to contain a logarithm and becomes $c\{(bp_1)^k - 1\}$ or $-c\{1 - (b - bp_1)^k\}$, where c is the initial capital of the expert.

The above would be a method of introducing piece-work into the Meteorological Office. The weather-forecaster would lose money whenever he made an incorrect forecast.

When making a probability estimate it may help to *imagine* that you are to be paid in accordance with the above scheme. (It is best to tabulate the amounts to be paid as a function of p_1 .)

9. Legal and Statistical Procedures Compared

In legal proceedings there are two men A and B known as lawyers and there is a hypothesis H . A is paid to pretend that he regards the probability of H as 1 and B is paid to pretend that he regards the probability of H as 0. Experiments are performed which consist in asking witnesses questions. A sequential procedure is adopted in which previous answers influence what further questions are asked and what further witnesses are called. (Sequential procedures are very common indeed in ordinary life.) But the jury which has to decide whether to accept or to reject

* For more than two alternatives the corresponding formula for b is $\log b = -\sum x_i \log x_i$, the initial "entropy".

H (or to remain undecided) does not control the experiments. The two lawyers correspond to two rival scientists with vested interests and the jury corresponds to the general scientific public. The decision of the jury depends on their estimates of the final probability of H . It also depends on their judgments of the utilities. They may therefore demand a higher threshold for the probability required in a murder case than in a case of petty theft. The law has never bothered to specify the thresholds numerically. In America a jury may be satisfied with a lower threshold for condemning a black man for the rape of a white woman than vice versa (*News-Chronicle*, 32815 (1951), 5). Such behaviour is unreasonable when combined with democratic bodies of decisions.

The importance of the jury's coming to a definite decision (even a wrong one) was recognized in law at the time of Edward III (c. 1350). At that time it was regarded as disgraceful for a jury not to be unanimous, and according to some reports such juries could be placed in a cart and upset in a ditch (*Enc. Brit.*, 11th ed., 15, 590). This can hardly be regarded as evidence that they believed in credibilities in those days. I say this because it was not officially recognized that juries could come to wrong decisions except through their stupidity or corruption.

10. *Minimax Solutions*

For completeness it would be desirable now to expound Wald's theory of statistical decision functions as far as his definition of Bayes solutions and minimax solutions. He gets as far as these definitions in the first 18 pages of *Statistical Decision Functions*, but not without introducing over 30 essential symbols and 20 verbal definitions. Fortunately it is possible to generalize and simplify the definitions of Bayes and minimax solutions with very little loss of rigour.

A number of mutually exclusive statistical hypotheses are specified (one of them being true). A number of possible decisions are also specified as allowable. An example of a decision is that a particular hypothesis (or perhaps a disjunction of hypotheses) is to be acted upon without further experiments. Such a decision is called a "terminal decision". Sequential decisions are also allowed. A sequential decision is a decision to perform further particular experiments. I do not think that it counts as an allowable decision to specify the final probabilities of the hypotheses, or their expected utilities. (My use of the word "allowable" here has nothing to do with Wald's use of the word "admissible".) The terminal and sequential decisions may be called "non-randomized decisions". A "randomized decision" is a decision to draw lots in a specified way in order to decide what non-randomized decision to make.

Notice how close all this is to being a classification of the decisions made in ordinary life, i.e., you often choose between (i) making up your mind, (ii) getting further evidence, (iii) deliberately making a mental or physical toss-up between alternatives. I cannot think of any other type of decision. But if you are giving advice you can specify the relative merits or expected utilities of taking various decisions, and you can make probability estimates.

A non-randomized decision function is a (single-valued) function of the observed results, the values of the function being allowable non-randomized decisions. A randomized decision function, δ , is a function of the observed results, the values of the function being randomized decisions.

Minus the expected utility for a given statistical hypothesis F and a given decision function δ is called the *risk* associated with F and δ , $r(F, \delta)$. (This is intended to allow for utilities including the cost of experimentation. Wald does not allow for the cost of theorizing.) If a distribution ξ of initial probabilities of the statistical hypotheses is assumed, the expected value of $r(F, \delta)$ is called $r^*(\xi, \delta)$, and a decision function δ which minimizes $r^*(\xi, \delta)$ is called a *Bayes solution* relative to ξ .

A decision function δ is said to be a *minimax solution* if it minimizes $\max_F r(F, \delta)$.

An initial distribution ξ is said to be *least favourable* if it maximizes $\min_\delta r^*(\xi, \delta)$. Wald shows under weak conditions that a minimax solution is a Bayes solution relative to a least favourable initial distribution. Minimax solutions seem to assume that we are living in the worst of all possible worlds. Mr. R. B. Braithwaite suggests calling them "prudent" rather than "rational".

Wald does not (in his book) explicitly recommend the adoption of minimax solutions, but he considered their theory worth developing because of its importance in the theory of statistical decision functions as a whole. In fact the book is more concerned with the theory than with

recommendations as to how to apply the theory. There is, however, the apparently obvious negative recommendation that Bayes solutions cannot be applied when the initial distribution ξ is unknown. The word "unknown" is rather misleading here. In order to see this we consider the case of only two hypotheses H and H' . Then ξ can be replaced by the probability, p , of H . I assert that in most practical applications we regard p as bounded by inequalities something like $\cdot 1 < p < \cdot 8$. For if we did not think that $\cdot 1 < p$ we would not be prepared to accept H on a small amount of evidence. Is ξ unknown if $\cdot 1 < p < \cdot 8$? Is it unknown if $\cdot 4 < p < \cdot 5$; if $\cdot 499999 < p < \cdot 500001$? In each of these circumstances it would be reasonable to use the Bayes solution corresponding to a value of p selected arbitrarily within its allowable range.

In what circumstances is a minimax solution reasonable? I suggest that it is reasonable if and only if the least favourable initial distribution is reasonable according to your body of beliefs. In particular a minimax solution is always reasonable provided that only reasonable ξ 's are entertained. But then the minimax solution is only one of a number of reasonable solutions namely all the Bayes solutions corresponding to the various ξ 's.

It is possible to generalise Wald's theory by introducing a distribution ζ of the ξ 's themselves. We would then be using a probability type-chain of type 3. (See section 3.) The expected value of $r^*(\xi, \delta)$ could be called $r^{**}(\zeta, \delta)$, and a decision function δ which minimizes $r^{**}(\zeta, \delta)$ could be called a "Bayes solution of type 3" relative to ζ . For consistency we could then call Wald's Bayes solutions "Bayes solutions of type 2". When there is only one available statistical hypothesis F we may define a "Bayes solution of type 1" (relative to F) as one which minimizes $r(F, \delta)$. The use of Bayes solutions of any type is an application of the principle of rational behaviour.

One purpose in introducing Bayes solutions of the third type is in order to overcome feelings of uneasiness in connection with examples like the one mentioned above, where $\cdot 1 < p < \cdot 8$. One feels that if $p = \cdot 09$ has been completely ruled out, then $p = \cdot 11$ should be *nearly* ruled out, and this can only be done by using probability type-chains of the third type. It may be objected that the higher the type the woollier the probabilities. It will be found, however, that the higher the type the less the woolliness matters, provided the calculations do not become too complicated. Naturally any lack of definition in ζ is reflected in ambiguity of the Bayes solution of type 3. This ambiguity can be avoided by introducing a *minimax solution of type 2*, i.e., a decision function which minimizes $\max_{\xi} r^*(\xi, \delta)$.

By the time we had gone as far as type-chains of type 3 I do not think we would be inclined to objectify the degrees of belief any further. It would therefore probably not be necessary to introduce Bayes solutions of type 4 and minimax solutions of type 3.

Minimax solutions (of type 1) were in effect originated by von Neumann in the theory of games, and it is in this theory that they are most justifiable. But even here in practice you would prefer to maximize your expected gain. You would probably use minimax solutions when you had a fair degree of belief that your opponent was a good player. Even when you use the minimax solution you may be maximizing your expected gain, since you may already have worked out the details of the minimax solution, and you would probably not have time to work out anything better once a game had started. To attempt to use a method other than the minimax method would then lead to too large a probability of a large loss, especially in a game like poker. (As a matter of fact I do not think the minimax solution has been worked out for poker.)

I am much indebted to the referee for his critical comments.