When simple alternatives to Bayes formula work well: Reducing the cognitive load when updating probability forecasts

Paul Goodwin

23 June 2014

The Management School, University of Bath, Bath, BA2 7AY,United Kingdom

Email: mnspg@bath.ac.uk.

Acknowledgements

Kesten Green, Scott Armstrong, Konstantinos Katsikopoulos and two anonymous reviewers provided helpful comments enabling the author to make improvements to the paper. Gerd Gigerenzer provided very useful suggestions for improving the section on natural frequencies.

The author confirms that he has read each of the original studies cited. Where there was any doubt as to the interpretation of a cited paper he attempted to contact the authors to confirm the interpretation and to ascertain whether any other publications should be cited.

When simple alternatives to Bayes formula work well: Reducing the cognitive load when updating probability forecasts

Abstract

Bayes theorem is the normative method for revising probability forecasts when new information is received. However, for unaided forecasters its application can be difficult, effortful, opaque and even counter-intuitive. Two simple heuristics are proposed for approximating Bayes formula while yielding accurate decisions. Their performance was assessed: i) where a decision is made on which of two events is most probable and ii) where a choice is made between an option yielding an intermediate utility for certain or a gamble which will result in either a worse or better utility ('certainty or risk' decisions). For 'most probable event' decisions the first heuristic always results in the correct decision when the reliability of the new information does not depend on which event will occur. In other cases the second heuristic typically led to the correct decision for about 95% of 'most probable event' decisions and 86% of 'certainty or risk' decisions.

Keywords: Bayes theorem, forecasting, heuristics, probability estimation

When simple alternatives to Bayes formula work well: Reducing the cognitive load when updating probability forecasts

1.1 Introduction

Forecasts are often expressed as probabilities and, when new information is received, Bayes theorem provides the normative way of revising these prior probabilities. For example, economic forecasters may revise their subjective probabilities of a recession upwards if an economic leading indicator suggests that a decline in growth is on the horizon (Schnader and Stekler 1998). Similarly, estimates of the probability of success of a potential new product may be revised upwards when encouraging market research results become available. Many situations involve estimating probabilities for two mutually exclusive and exhaustive events, A and $\overline{A}$ (e.g., recession or no recession or rain or no rain). In this case Bayes theorem can be stated as:

$$P(A|N) = \frac{Po(A) \times P(N|A)}{Po(A) \times P(N|A) + (1-Po(A)) \times P(N|\overline{A})} \qquad (1)$$

Where: Po(A) is the prior probability of event A

P(A|N) is the posterior probability of A

$\overline{A}$ is the event which is complementary to A

N = the new information

P(N|A) is the likelihood of the new information, or the probability of getting the new information given that A will occur

Where $p(N|A) = 1 - p(N|\overline{A})$ the new information will be referred to as a 'symmetric indicator'. For example, suppose that a test will indicate whether a manufactured component is defective or non-defective. If it has the same probability of giving a correct indication irrespective of whether or not a component is defective then it will provide a symmetric indication of the component's condition. If this condition does not apply then the new information is an 'asymmetric indicator'.

Applying Bayes theorem can pose difficulties for unaided forecasters. The cognitive effort involved in using the formula may make it unacceptable when a quick decision needs to be made or when a calculator is unavailable. To those unfamiliar with probability theory the formula may lack transparency and hence there may be a distrust of the posterior probability produced by it. For example, suggestions that jurors should use Bayes theorem to determine the probability of a defendant's guilt have never been widely implemented because of the difficulties involved in getting people to apply it, or accept it, even when they are provided with a structured format (Balding 1997). There is also plenty of evidence that, in many circumstances, people do not naturally revise probabilities according to Bayes theorem.

These factors suggest that it may be worth trying to develop simple, but reliable, approximations to Bayes theorem, which people could be encouraged to employ when use of the exact formula is impractical. This paper therefore addresses two questions.

1. Is it possible to identify simple and intuitively appealing heuristics that will approximate Bayes theorem and lead to the same decisions in a wide range of circumstances?

2. Under what conditions, if any, would the use of these heuristics lead to serious errors?

2.1 Background

Some researchers have assumed in their models of human prediction and decision making that people revise their prior beliefs according to (1) (e.g., Schnader and Stekler 1998). However, a substantial body of research has found that in many circumstances this is not the case. A predominant finding of the literature of the 1960s was that people are conservative in that they insufficiently revise their prior probabilities when they receive diagnostic new information when compared to the revisions prescribed by the theorem (e.g., Phillips and Edwards 1966; Phillips et al. 1966; Edwards 1968). Later work has suggested the opposite in that people underweight prior probabilities and make their judgment primarily on how representative the new information appears to be of either A or $\overline{A}$ (Grether 1992; Mahmoud and Grether 1995; Charness, Karni and Levin, 2007; Holt and Smith 2009). The difference between these findings may, in part, reflect whether the prior probabilities were estimated by the forecaster themselves or whether they were supplied to them (other factors, like incentives for accurate judgment, may also have played a role). A self-estimated prior probability would be likely to carry greater salience and hence greater weight in the revision process than a supplied probability (e.g. Phillips and Edwards 1966; Evans, Handley and Over 2002). Indeed, it may act as an anchor in an "anchor and adjustment" process (Tversky and Kahneman 1974).

Recent research by Goodwin et al. (2013) suggests a simple model can represent people's revisions to their *own* prior probabilities. This model is a weighted average of the prior probability and the likelihood associated with the new information when A occurs:

$$P(A|N) = 0.66\ Po(A) + 0.41\ P(N|A) \qquad (2)$$

In all cases examined by Goodwin et al. (2013) the new information was a symmetric indicator. The estimated weights of 0.66 and 0.41 were obtain by applying generalized estimating equations to the prior and posterior probability estimates of 54 participants in an experiment. Each participant judged the probability of a recession in nine scenarios both before and after receiving information from an economic indicator.

Barbey and Sloman (2007) discuss a number of theoretical accounts of how Bayesian estimation can be facilitated. There is strong evidence that people are likely to revise their prior probabilities more accurately when information is presented in a natural frequency, rather than a probability, format (Gigerenzer and Hoffrage 1995; Cosmides and Tooby 1996; Koehler 1996; Gigerenzer, Gaissmaier, Kurz-Milcke, Schwartz and Woloshin 2007). Goodwin and Wright (1991) demonstrated this method and it was discussed by Kleiter (1992). Gigerenzer (2011) defines a natural frequency as a joint frequency of two events. For example, it could be the number of components manufactured in a factory that are both defective *and* have been found to be defective in a quality control test that is not perfectly reliable. Figure 1 shows how 1000 typical components could lead to four natural frequencies depending on whether or not they are defective and whether or not they have failed the test (the natural frequencies are at the bottom of the tree). The probability that a component is defective, given that it has failed the test, can be easily determined from the diagram. Of the 140 components that failed the test, 110 are defective. Hence the required probability is simply calculated as 110/140 = 78.6%.

Figure 1 here

Why do natural frequencies make Bayesian revision easier? One theory is that humans have evolved a component of their brain that efficiently processes natural frequency information (Gigerenzer 2000) (Chapter 4). The frequency with which events are experienced has been the natural way in which humans (and even animals) have obtained information on the risks they face throughout their evolution, while the use of probabilities is relatively recent (Gigerenzer 2002). In addition, the natural frequency format may allow decision makers to have an accurate perception of the set structure underlying the necessary calculation (see Evans, Handley, Perham, Over and Thompson 2000; Barbey and Sloman 2007). The use of Euler diagrams to represent the set structure has also been found to aid Bayesian inference (Sloman, Over, Slovak and Stibel 2003).

However, in some circumstances in may be difficult or unnatural to conceive the problem in terms of a large population of repeated similar events. For example, consider the task of estimating the probability that an innovative new product will make a profit, given that market research has indicated that it will, or the probability that a specific construction project will be delayed given that geological tests have indicated problems with the local rock structure. On other occasions decision makers may not have the time, commitment or even the need to estimate perfectly accurate Bayesian posterior probabilities, particularly as decision problems often have a wide degree of tolerance to errors in the underlying probability estimates (von Winterfeldt and Edwards 1986). Also, the importance of getting the decision right may not be regarded as crucial, so the need to avoid an erroneous choice does not justify the effort required to estimate correct probabilities (Payne, Bettman and Johnson 1993).

Indeed, in some circumstances where judgments are applied to decisions greater effort may even lead to less accurate judgments (e.g., see Katsikopoulos (2011) for a review). All of this suggests that there may be a useful role for heuristics that can handle information in the form of probabilities and yet be cognitively less demanding than Bayes formula. Ideally, the heuristics should be intuitively reasonable and carry an acceptably low risk of yielding the wrong decision.

When the heuristic (2) reported in Goodwin, et al. (2013) was applied to a range of decisions, discrepancies between decisions based on probabilities revised according to Bayes theorem and those based on the heuristic were relatively rare. Moreover, when they occurred they were generally inconsequential in that the differences between the expected utilities of the decision based on Bayes theorem and those of the discrepant decision were small. This finding raises the possibility of being able to recommend to decision makers a simple rule or rules that will accord with their natural way of thinking and will give them a high probability of making a correct decision. This possibility is explored next in the context of two types of decision.

2.1 Deciding which event is most probable

On many occasions people have to decide which of two mutually exclusive and exhaustive events is most likely to occur. Is it more probable that the price of a stock will rise rather than fall over the next month? Is economic growth over the next three years more probable than a decline in GDP? Is a default by a debtor more probable than no default? Is it more probable that tomorrow will be a rain-free day or a day when some precipitation will occur?

This choice is shown in the simple decision tree in figure 2 where a forecast of A, F(A), would be chosen if event A is most probable and a forecast of $\overline{A}$ (i.e., F($\overline{A}$))

would be chosen otherwise. The probability of A is w and the 0's and 1's are the decision maker's utilities for the worst and best outcomes respectively (assume that utilities are measured on a 0 to 1 scale throughout this paper). These utilities assume that correctly forecasting the event is equally good irrespective of which option is chosen. For example, it assumes that the forecaster will be just as satisfied with a correct forecast of rain and a correct forecast of fine weather. Similarly, choosing the wrong option is assumed to be equally bad, irrespective of which event actually occurs. When new information is received let $w = p_B$ if the Bayesian posterior probability is used to make the decision and $w = p_E$ if the decision maker's estimate of the posterior probability is used. The decision will differ only if $p_E > 0.5$ when $p_B < 0.5$ or when $p_E < 0.5$ when $p_B > 0.5$, that is when the decision maker's probability and the Bayesian posterior probability are on 'opposite' sides of 0.5.

Figure 2 here

When the decisions do differ how serious will this be? Expected utility loss is the expected loss in the decision maker's satisfaction caused by the discrepancy. It is the difference between the expected utility of the best option and the option selected, with both expected utilities calculated using the correct Bayes posteriors. For 'most probable event' decisions it is $|2p_B - 1|$. This is because if F(A) is the correct decision then the expected utility is p_B. An incorrect choice of $F(\overline{A})$ would yield an expected utility of $1 - p_B$ so the difference (or utility loss) is $p_B - (1 - p_B) = 2p_B - 1$. When $F(\overline{A})$ is the correct choice the utility loss is $1 - 2p_B$.

2.1.1 Symmetric indicator

Consider first situations where the new information is a symmetric indicator. The model in (2) suggests a very simple heuristic: *Take the Average*:

$$P(A|N) = 0.5\ Po(A) + 0.5\ P(N|A) \quad (3)$$

Here the revised probability is just the mean of the prior probability and the likelihood associated with the event in question. If the result exceeds 0.5 then A is considered to be the most probable event. For example, suppose that a forecaster estimates the prior probability of a recession in the next year, Po(A) to be 0.2. An economic forecast is then published which predicts that there will be a recession in the next year. Suppose also that the probability of the forecast predicting a recession given that there will be a recession, P(N|A), is 0.7. The simple heuristic yields an estimated posterior probability of a recession of 0.45. Since this probability is less than 0.5, a person using the heuristic would conclude that "no recession" is the more probable event. If the economic forecast is a symmetric indicator, Bayes theorem yields a posterior probability of 0.37 so a person using the theorem would agree that "no recession" is more probable.

How well would Take the Average work in general when applied to a 'most probable event' decision and when the indication is symmetric? In fact, it would give the same decision as Bayes theorem 100% of the time, as shown below. When the posterior probability P(A|N) is greater than 0.5, according to Bayes theorem:

$$\frac{Po(A)\ P(N|A)}{Po(A)\ P(N|A) + [1-Po(A)][1- P(N|A)]} > 0.5 \quad (4)$$

So: $$2\ Po(A)\ P(N|A) > Po(A)\ P(N|A) + [1-Po(A)]\ [1- P(N|A)] \quad (5)$$

This expression simplifies to:

$$Po(A) + P(N|A) > 1 \quad \text{or} \quad 0.5\ Po(A) + 0.5\ P(N|A) > 0.5 \quad (6)$$

2.1.2 Asymmetric indicator

Take the Average cannot be guaranteed to yield the correct decision when an indicator is asymmetric. This is because it is ignoring the value of $P(N|\overline{A})$ so it would give the same result irrespective of what this likelihood is. For example, consider the forecasting problem referred to above, where the prior probability of a recession is 0.2. Suppose that the economic forecast has a 0.7 probability of forecasting a recession when there will be a recession, but only a 0.05 probability of forecasting a recession when "no recession" will occur so that $P(N|A) = 0.7$ but $P(N|\overline{A}) = 0.05$. The heuristic's posterior probability of 0.45 will be on the 'opposite side' of 0.5 when compared to the Bayes posterior of 0.78.

In this case is it possible to derive an alternative heuristic which takes into account all of the information? When an indicator is asymmetric a perfectly correct decision can be guaranteed if the following procedure is followed.

1. Divide $P(N|\overline{A})$ by the sum of the likelihoods
2. Choose A as being most probable only if the prior probability Po(A) exceeds this ratio.

In the last version of the recession forecasting problem this procedure would result in 0.05/0.75 = 0.07. Hence a recession would be considered to be the most probable event as Po(A) = 0.2 so the decision would agree with that based on Bayes formula. This works because, when the posterior probability, P(A|N), exceeds 0.5.

$$Po(A)\,P(N|A) > [1-Po(A)]\,P(N|\overline{A}) \quad (7)$$

$$\text{so: } Po(A) > P(N|\overline{A})/[P(N|A) + P(N|\overline{A})] \quad (8)$$

Although these steps are guaranteed to give the correct decision and involve less effort that the application of the Bayes formula they may still be too complex for an unaided decision maker. If this is the case a simpler heuristic, *Sum the Pros, Sum the Cons,* will often give good approximate results as shown below.

Note that the left hand side of the inequality in (4) is the product of the probabilities 'favouring' A. For example the second term is the probability of obtaining the new information if A will occur. The right hand side is the product of the probabilities 'disfavouring' A. The comparison in (4) will be easier if sums replace the products, that is if the sum of 'favouring' probabilities (the 'pros') exceed the 'disfavouring' probabilities (the 'cons'). This approximation should work well because if:

$$a.b>(1-a)c$$

it is likely that $a + b > (1-a) + c$ when $0 \leq a,b,c \leq 1.0$.

Indeed this proved to be the case in over 96% of cases when it was tested over all combinations of a, b and c with the values varying in steps of 0.01 (Similarly when $a + b > (1-a) + c$ then $a.b>(1-a)c$ in over 94% of cases). Thus the Sum the Pros, Sum the Cons heuristic is: assume A is more probable than $\overline{A}$ if :

$$Po(A) + P(N|A) > [1-Po(A)] + P(N|\overline{A})] \quad (9)$$

For the last version of the recession forecasting example the left hand side is 0.9 (0.2+0.7) and a right hand side of 0.85 (0.8 + 0.05) so the heuristic agrees with the Bayes posterior that a recession is more probable than no recession. It can be seen that Sum the Pros, Sum the Cons is a generalisation of Take the Average as, if the indication from the new information is symmetric, the sum of all four probabilities in

(6) will be two so the left hand side will have an average exceeding 0.5 when A is the more probable event.

Sum the Pros, Sum the Cons was tested for 'most probable event' decisions across the following ranges of values (these will be referred to as the test set for most probable event decisions):

- prior probabilities, Po(A), from 0 to 1 in steps of 0.01
- values of P(N|A) from 0.01 to 0.99 in steps of 0.01
- values of $P(N|\overline{A})$ from 0.01 and 0.99 in steps of 0.01.

Figure 3 shows the percentage of times the heuristic led to a discrepancy with decisions based on Bayes for different prior probabilities.

Figure 3 here

The heuristic performed at its worst if the prior probability is 0.3 or 0.7, when about 8.5% of decisions were discrepant (i.e., about 91.5% of decisions were still correct). Overall, only 5.5% of decisions differed from those based on Bayes theorem (i.e., 94.5% agreed). Note that the disagreements are minimised when the prior probability is at the extremes or close to 0.5. Extreme priors will tend to lead to posterior probabilities which are either well below or well above 0.5 using both Bayes formula and the heuristic so there will usually be no disagreement between the two on which is the most probably event. If the prior probability for event A is 0.5 then, according to Bayes rule, A will be more probable than $\overline{A}$ if $P(N|A) > P(N|\overline{A})$. Under these circumstances, Sum the Pros, Sum the Cons is always bound to agree with Bayes rule because it would indicate that A is more probable if $0.5 + P(N|A) > 0.5 + P(N|\overline{A})$ which will only be true if $P(N|A) > P(N|\overline{A})$.

These results suggest that the risk of wrongly identifying the most probable event when using this heuristic is low. But how serious are the discrepancies? When discrepancies occurred the expected utility loss was, on average, 0.21 indicating that typically the loss of satisfaction by the decision maker was 21% of the difference between the utilities of the worst and best possible outcomes. However, given that discrepancies were rare, over all decisions the expected utility loss was only 0.01.

2.2 Choosing between a certain or risky alternative

Figure 4 displays a decision tree for a second commonly encountered type of decision problem. Here the decision maker has to choose between a 'risk free' option and a gamble which will result in either a worse or better outcome than the risk free option. Specifically the decision involves two alternative courses of action D_1 and D_2. Alternative D_1 can result in two outcomes A and $\overline{A}$. The utilities that can be obtained are shown at the ends of the branches and range from 0 (the worst outcome) to 1 the best with $0 \leq U \leq 1$. Note that D_2 always leads to a utility of U. The probability of the best outcome, A , is x.

Figure 4 here

If the decision maker receives new information relating to the probability of outcome A then let $x = p_B$ if the Bayesian posterior probability is used to make the decision and $x = p_E$ if the decision maker's estimate of the posterior probability is used. According to the axioms of utility theory (e.g., see Goodwin and Wright (2014)) the decision maker will be indifferent between D_1 and D_2 when $U = x$. Discrepant decisions will therefore be made if: $p_E > U$ when $p_B < U$ or vice versa, that is when p_E and p_B are on 'opposite sides' of U. The expected utility loss of a discrepant decision will be the difference between expected utilities of the correct and incorrect decisions based on the Bayes probabilities, that is: $|U - p_B|$. For example, if D_1 is the correct decision and D_2 is chosen then the utility loss will be: $1.p_B +0.(1- p_B) –U = p_B –U$.

2.2.1 Symmetric indicator

The performance of Take the Average was tested on 'certainty or risk' decisions when the indication from the new information was symmetric. This testing was done for all combinations of values of Po(A) and p(N|A) from 0.01 to 0.99 in steps of 0.01 when U =0.1, 0.3, 0.5, 0.7 and 0.9. The estimate of the posterior probability was simply: 0.5 [(Po(A) + p(N|A)]. Over all values the heuristic yielded a decision that agreed with the one based on Bayes theorem on 86.3% of occasions (see figure 5) and the mean expected utility loss was only 0.008. When U =0.5 there was 100% agreement between the Bayes decision and the heuristic. Less than 70% agreement occurred where both U and Po(A) were low (e.g., $U<0.1$ and $Po(A)<0.2$) or both high (e.g., $U>0.9$ and $Po(A)>0.8$) with the agreement percentage getting worse as the pairs of values became more extreme. In these situations the expected utilities of A and $\bar{A}$ will tend to be close so there is a greater chance of a discrepancy, though the expected utility loss of any discrepancy will be low. On average the mean expected utility loss when discrepancies occurred was only 0.085. It is interesting to note the weighted average heuristic (2) typically adopted by participants in the study by Goodwin et al. (2013), which is similar to Take the Average, was well adapted to the task.

2.2.2 Asymmetric indicator

The performance of Sum the Pros, Sum the Cons was tested on 'certainty or risk' decisions when the indication from the new information was asymmetric. In this case the posterior probability was obtained as:

$$P(A|N) = \frac{Po(A) + P(N|A)}{Po(A) + P(N|A) + [1-Po(A)] + P(N|\bar{A})]} \quad (10)$$

This is Bayes formula (1) with sums replacing products. It may look complex for a heuristic but it is simply:

$$\frac{\text{Sum of the Pros}}{\text{Sum of the Pros and Cons}}$$

When the indication is symmetric it simplifies to Take the Average. The heuristic was tested on 'certainty or risk' decisions for all combinations of values of Po(A), p(N|A) and $P(N|\bar{A})$ from 0.01 to 0.99 in steps of 0.01 when U =0.1, 0.3, 0.5, 0.7 and 0.9. Combinations where $P(N|A) = 1 - P(N|\bar{A})$ were excluded. The performance was very similar to that when the indicator was symmetric. Over all values the heuristic yielded a decision that agreed with the one based on Bayes theorem on 85.6% of occasions and the mean expected utility loss was only 0.013. Figure 5 shows the percentage of occasions when the heuristic disagreed with the Bayes decision for the different values of Po(A). As before, levels of agreement were below 70% when both U and Po(A) were low (e.g., $U<0.1$ and $Po(A)<0.15$) or both were high (e.g., $U>0.9$ and $Po(A)>0.85$). When there were discrepancies the mean expected utility loss was 0.091

Figure 5 here

3. Discussion and Conclusions

Bayes theorem is the normative method for revising probabilities when new information is received so the heuristics cannot surpass it in terms of accuracy. However, they can be superior in terms of the cognitive effort that they require and in their acceptability to decision makers. Decision makers seek to balance the cognitive effort they put into their decisions against the desire to maximise the chances of making the correct choice (Payne et al. 1993). If a heuristic has a high probability of leading to an accurate choice and involves relatively little cognitive effort it is likely to be acceptable. However, acceptability may also depend on the intuitive reasonableness of the heuristic -does it appear to make sense? For example, it is known that company sales forecasters tend to make too many judgmental adjustments to the statistical forecasts generated by their computer systems (Fildes, Goodwin, Lawrence and Nikolopoulos 2009). Yet the least effortful strategy would be merely to accept these forecasts without change. It seems that many of the adjustments are made because the forecasters do not understand the algorithms that have generated the statistical forecasts or their rationale. They regard them as a 'black box'. In particular, they perceive patterns in the random movements in sales time series and see the computer system's discounting of these movements as lacking intuitive reasonableness. As a result the forecasts are deemed to be unacceptable and they are changed. The heuristics presented above do appear to meet the three requirements of i) requiring less mental effort to implement than Bayes theorem, ii) providing a high chance of yielding an accurate choice, and iii) being intuitively reasonable.

Decisions often have what von Winterfeldt and Edwards (1986) refer to as 'flat maxima', that is the optimum choice is relatively insensitive to errors in the estimates of probabilities and utilities. The heuristics are able to exploit this property. This tolerance is particularly evident for 'most probable event' decisions. In particular, when the indicator is symmetric it is just not worth going to the effort of applying the Bayes formula. Take the average will guarantee perfect accuracy. In other cases the use of the heuristics clearly would not be advisable. This will be the case in a 'certainty or risk' decision when both the prior probability, Po(A), and the utility, U are both very low or both very high. For example, when Po(A) = 0.02 and U = 0.1 Sum the Pros, Sum the Cons has a 77% probability of indicating the wrong choice (when averaged over all values of P(N|A) and p(N|$\overline{A}$)). As a rule of thumb, if *both* values are below 0.2 or both are above 0.8 the heuristic should be avoided. The extent to which the heuristics should be used in between these extremes of good and bad performance is, of course a judgment call. In many situations a heuristic providing an 85% probability of an accurate decision is likely to be acceptable, given the reduced cognitive effort involved. It will clearly not be when decisions are of high importance (e.g., life and death decisions).

Interestingly, this notion of using simple rules, based on averages and sums, to update prior estimates was suggested in the 1950's in the context of regression analysis. Armstrong (1985) describes an approach which he terms the 'poor man's Bayesian regression analysis'. It involves a first step where *a priori* estimates of a model's coefficients are averaged with those estimated using regression analysis (e.g., using least squares). When Tessier and Armstong (2014) applied the approach to sales estimation in the US lodging market they found that it improved the accuracy of the estimates, while at the same time incurring little cost.

The analysis presented in this paper has a number of limitations. It was assumed that the new information, N, was received correctly, and more importantly, that P(N|A) and P(N|$\overline{A}$) are also known correctly. Nevertheless, there are many circumstances where P(N|A) and P(N|$\overline{A}$) are likely to be known. For example, the accuracy of medical tests or electronic tests used in quality control is often known, while information on the accuracy of weather forecasts is widely available. Where these values are not known exactly an error in their estimation would apply equally to Bayes formula and to the heuristics so their relative accuracy would remain unchanged. Nevertheless, there is potentially scope for the development of heuristics to support the estimation of likelihoods. Secondly, for a given prior probability Po(A), when the indication from the new information was asymmetric, the results assumed that all combinations of values of P(N|A) and P(N| $\overline{A}$) between 0.01 and 0.99 were equally likely to apply (i.e., a bivariate uniform distribution was assumed). In practical problems particular combinations of these values may be more common, but it is of course, difficult, if not impossible, to establish this. In addition, the research has only considered decisions with two options and up to two discrete outcomes, though these types of decisions *are* likely to be commonly encountered in practice.

Probability forecasts are important in many practical contexts. Taken together, the results provide strong evidence that, when these forecasts involve revisions based on new information, simpler can often be best.

References

Armstrong J. S. Long range Forecasting. From Crystal Ball to Computer. New York: John Wiley, 1985.

Balding D. Probable cause. The Times Higher 1997; October 24: 23.

Barbey A. K., Sloman S. A. Base-rate respect: From ecological rationality to dual processes. Behavioral and Brain Sciences 2007; 30 (3): 241-+.

Charness G., Karni E., Levin D. Individual and group decision making under risk: An experimental study of Bayesian updating and violations of first-order stochastic dominance. Journal of Risk and Uncertainty 2007; 35 (2): 129-148.

Cosmides L., Tooby J. Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition 1996; 58 (1): 1-73.

Edwards W. Conservatism in human information processing. In: B.Kleinmuntz editor. Formal Representation of Human Judgment. New York: Wiley, 1968. pp. 17-52.

Evans J. St., Handley S. J., Over D. E., Perham N. Background beliefs in Bayesian inference. Memory & Cognition 2002; 30 (2): 179-190.

Evans J. St., Handley S. J., Perham N., Over D. E., Thompson V. A. Frequency versus probability formats in statistical word problems. Cognition 2000; 77 (3): 197-213.

Fildes R., Goodwin P., Lawrence M., Nikolopoulos K. Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. International Journal of Forecasting 2009; 25 (1): 3-23.

Gigerenzer G. Adaptive thinking: rationality in the real world. New York: Oxford University Press, 2000.

Gigerenzer G. Reckoning with risk. London: Penguin Books, 2002.

Gigerenzer G. The art of risk communication. What are natural frequencies? British Medical Journal 2011; 343.

Gigerenzer G., Gaissmaier W., Kurz-Milcke E., Schwartz L.M., Woloshin S.W. Helping doctors and patients make sense of health statistics. Psychological Science in the Public Interest 2007; 8: 53-96.

Gigerenzer G., Hoffrage U. How to improve Bayesian reasoning without instruction - frequency formats. Psychological Review 1995; 102 (4): 684-704.

Goodwin P. , Önkal D., Stekler H. O. Does it matter if decision makers' probability assessments do not conform to Bayes' theorem when they are facing risky decisions? University of Bath, School of Management, Working Paper 2013.

Goodwin P., Wright G. Decision Analysis for Management Judgment, 1st edition. Chichester, UK: John Wiley, 1991 pp166-167.

Goodwin P., Wright G. Decision Analysis for Management Judgment, 5th edition. Chichester, UK: John Wiley, 2014.

Grether D.M. Testing Bayes rule and the representativeness heuristic: Some experimental evidence. Journal of Economic Behavior and Organization 1992; 17: 31-57.

Holt C. A., Smith A. M. An update on Bayesian updating. Journal of Economic Behavior & Organization 2009; 69 (2): 125-134.

Katsikopoulos K. V. Psychological Heuristics for Making Inferences: Definition, Performance, and the Emerging Theory and Practice. Decision Analysis 2011; 8 (1): 10-29.

Kleiter G.D. Natural sampling: rationality without base rates. In: G.H Fischer, G. Laming editors. Contributions to Mathemetical Psychology, Psychometrics and Methodology. New York: Springer-Verlag 1994. pp. 375-388.

Koehler J. J. The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. Behavioral and Brain Sciences 1996; 19 (1): 1-&.

Mahmoud A.A., Grether D.M. Are people Bayesian? Uncovering behavioral strategies. Journal of the American Statistical Association 1995; 90: 1137-1145.

Payne J. W., Bettman J. R., Johnson E. J. The Adaptive Decision Maker. Cambridge, UK: Cambridge University Press, 1993.

Phillips L. D., Edwards W. Conservatism in a simple probability inference task. Journal of Experimental Psychology 1966; 72: 346-367.

Phillips L. D., Hays W.L., Edwards W. Conservatism in Complex Probabilistic Inference. IEEE Transactions on Human Factors in Electronics 1966; 7: 7-18.

Schnader M. H., Stekler H. O. Sources of turning point forecast errors. Applied Economics Letters 1998; 5 (8): 519-521.

Sloman S. A., Over D., Slovak L., Stibel J. M. Frequency illusions and other fallacies. Organizational Behavior and Human Decision Processes 2003; 91 (2): 296-309.

Tessier T.H., Armstong J.S. Decomposition of time-series forecasts by current level and change: effects on accuracy. *Journal of Business Research* 2014; (this issue).

Tversky A., Kahneman D. Judgment under uncertainty: Heuristics and biases. Science 1974; 185: 1124-1131.

von Winterfeldt D., Edwards W. Decision Analysis and Behavioral Research. Cambridge, UK: Cambridge University Press, 1986.

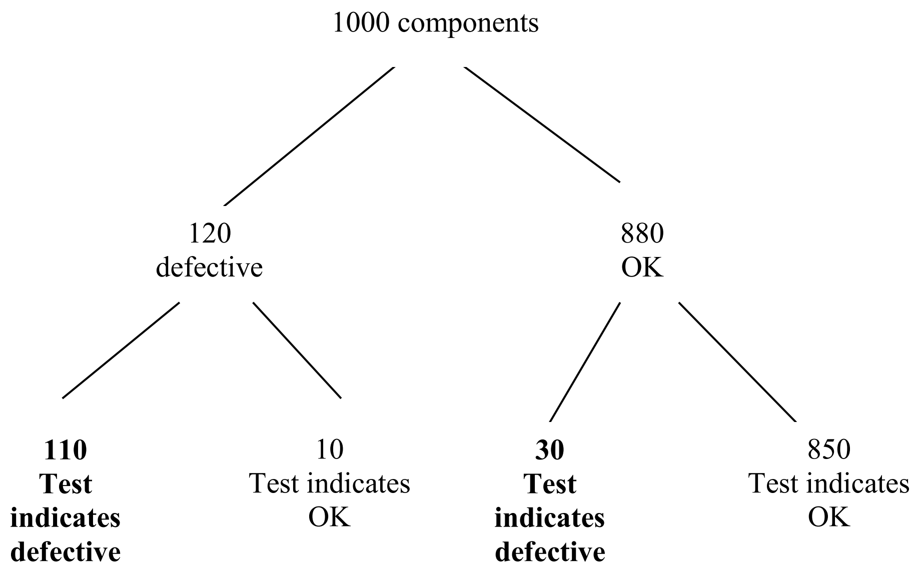

Figure 1. Using natural frequencies to estimate posterior probabilities

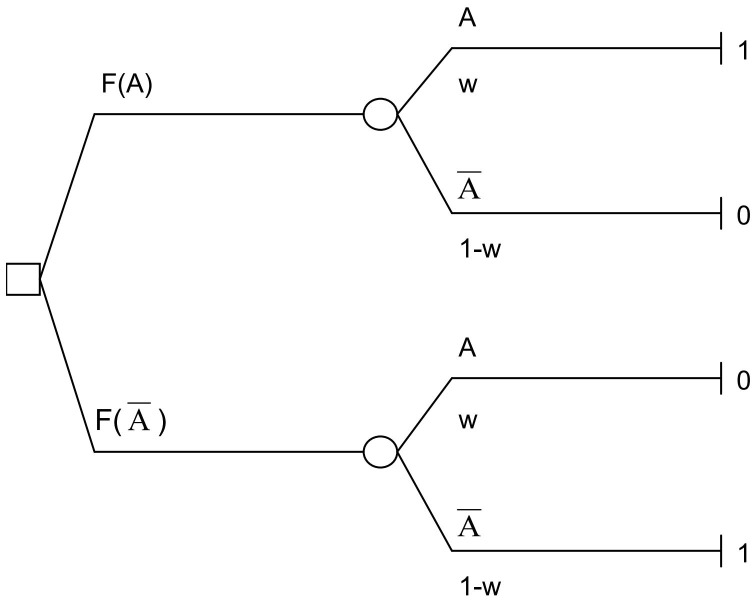

Figure 2 Decision tree for identifying the most probable event

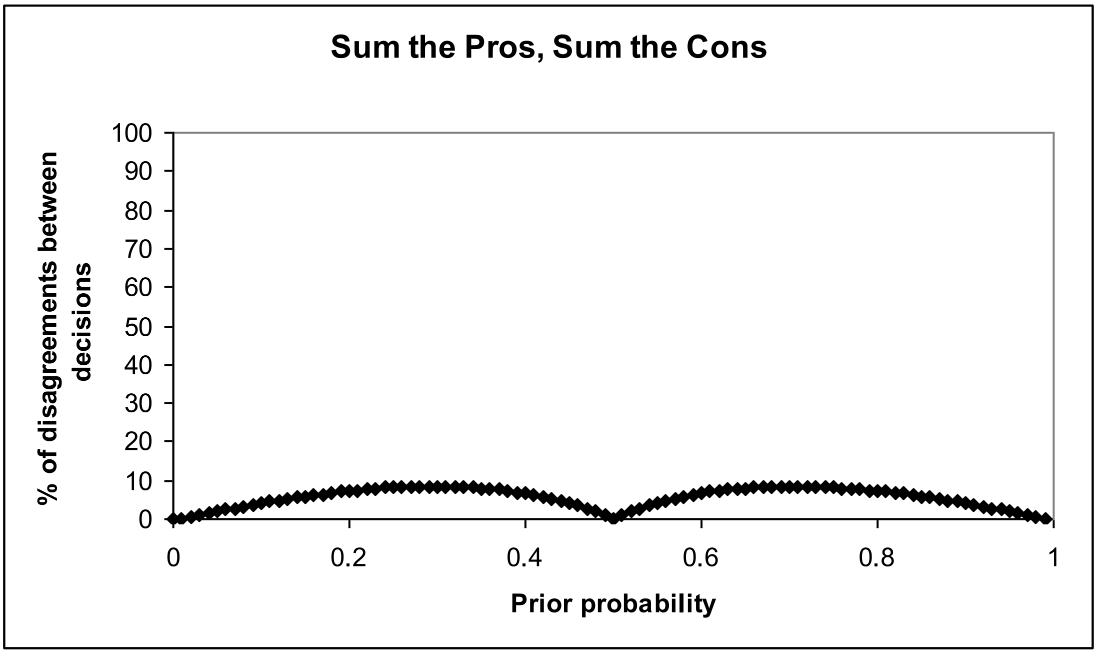

Figure 3 The performance of Sum the Pros, Sum the Cons when the indicator is asymmetric

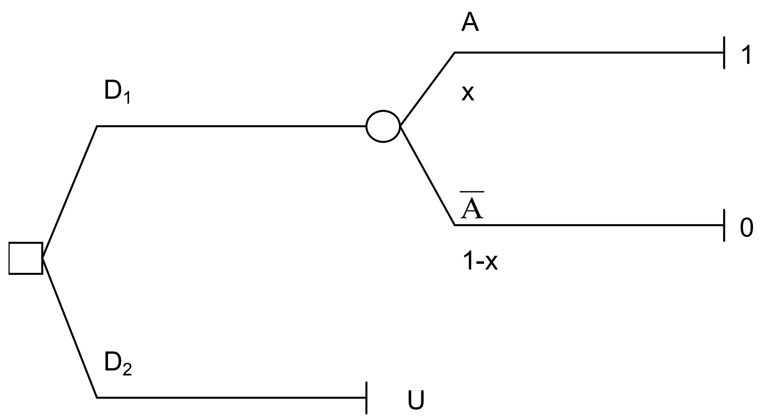

Figure 4. Decision tree for choosing between options with certain and risky outcomes where D_2 always leads to a utility of U.

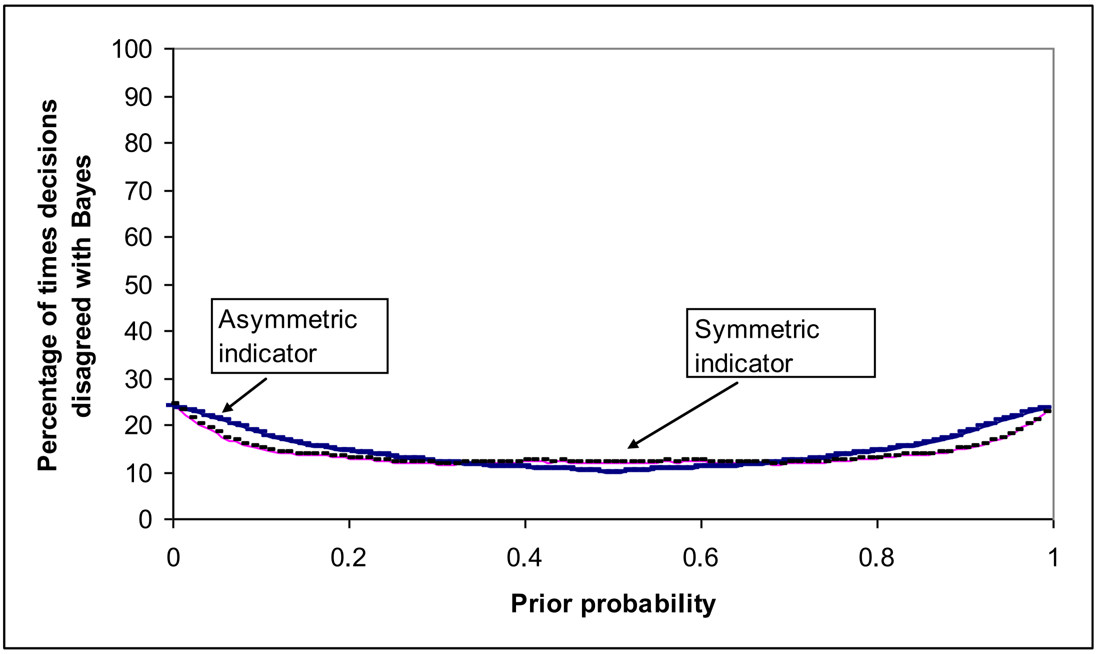

Figure 5 The performance of *Take the Average* and *Sum the Pros, Sum the Cons* for ‘certainty or risk’ decisions