

Ambiguity, the Certainty Illusion, and Gigerenzer's Natural Frequency Approach to Reasoning with Inverse Probabilities*

John Fountain[†] and Philip Gunby[‡]

2 February 2010
Version: 06 April 2010

Abstract

People have difficulty reasoning with information to do with uncertain situations, including when making economic decisions. This is especially true when decisions require the calculation of conditional probabilities. Putting the data in terms of natural frequencies as promoted by Gigerenzer (2002) makes it easier for people to reason in situations of uncertainty. Unfortunately, it invokes the normally false assumption that the frequency information is precise. The use of simple graphical techniques can help to resolve this problem, providing a tool that can be used for making decisions in uncertain situations when the probabilistic information is imprecise and thus ambiguity exists.

Keywords: Ambiguity; certainty illusion; inverse probability; natural frequencies; uncertainty.

JEL Classification: A200, D100, D800.

* Draft. The opinions and conclusions expressed are solely those of the authors. All errors are our own.

[†] Department of Economics, University of Canterbury.

[‡] Department of Economics, University of Canterbury.

[#] Corresponding Author: [REDACTED], ph: [REDACTED], fax: [REDACTED].

Ambiguity, the Certainty Illusion, and Gigerenzer's Natural Frequency Approach to Reasoning with Inverse Probabilities

Abstract

People have difficulty reasoning with information to do with uncertain situations, including when making economic decisions. This is especially true when decisions require the calculation of conditional probabilities. Putting the data in terms of natural frequencies as promoted by Gigerenzer (2002) makes it easier for people to reason in situations of uncertainty. Unfortunately, it invokes the normally false assumption that the frequency information is precise. The use of simple graphical techniques can help to resolve this problem, providing a tool that can be used for making decisions in uncertain situations when the probabilistic information is imprecise and thus ambiguity exists.

Keywords: Ambiguity; certainty illusion; inverse probability; natural frequencies; uncertainty.

JEL Classification: A200, D100, D800.

1. Introduction

The world is an uncertain place and many decisions people make depend on accurate information about the risks they face as well as clear thinking with the information available to them. If a medical treatment has the potential to improve a person's life but also potential side-effects which could worsen their life, should they go ahead with it? Medical tests providing information used in making this decision are prone to both false negatives and false positives, how much faith should a person place in a test result? When facing a choice about what risky financial assets to put their savings into, where the returns are uncertain and the capital investment is at risk, which assets should people choose? As with medical decisions, informative but imperfect signals are available, such as the credit ratings for the financial assets or the organisations issuing them. Other equally important decisions under uncertainty routinely occur such as when and how much to gamble, which party to vote for in an election, which consumption items to purchase, whether someone accused of a crime is guilty or innocent, and basing decisions on weather forecasts. These decisions all depend on clear thinking about probabilities.

Mistakes made in many of these situations can prove very costly. The Sally Clark case in the United Kingdom in which she was accused of killing her two children and was subsequently convicted, partly on the basis of incorrect calculation of probabilities by an expert witness, is an obvious example.* Whether or not a woman should undergo a mastectomy if a mammography is positive, or a man should have his prostate removed if his PSA test is positive are not as clear cut as it first seems. Both procedures involve risks of serious side-effects of physical discomfort, psychological stress, a perceived loss of femininity

* See *The Guardian*, (2007), 8 November, p. 11. The title of the article was "Sally Clark, Mother Wrongly Convicted of Killing Her Sons, Found Dead at Home." and her inability to recover from the accusations and initial conviction shows the potential cost of getting a probability calculation wrong.

for a mastectomy, and incontinence and impotence for removal of a prostate. Mistakes relating to how risky financial assets are can also prove very costly as seen with the recent financial crisis. For example, the California Public Employees Retirement System is reportedly suing three credit rating agencies for "...hundreds of millions in losses..." it made over what it saw as inaccurate credit ratings of financial assets. Irrespective of whether or not these assessments were accurate, they clearly affected beliefs about the riskiness of the assets. If people incorrectly adjust their beliefs because of poor statistical thinking, then serious consequences can result as is evident.

Thinking about how people make decisions in the face of risk is also very important to economic thought. The subjective expected utility framework is predicated on people calculating probabilities correctly when faced with decisions where they face uncertain situations. If people do not correctly reason about uncertainty then the validity of the current mainstream description of people's decision making under uncertainty is called into question. That possible problems may exist with the subjective expected utility model in economics as a descriptive summary of individuals' decisions under uncertainty has been around since the Allais and Ellsberg paradoxes were discovered. Both results showed that the observed choices of people conflict with the predictions of expected utility theory, although for different reasons. This resulted in generalisations of expected utility theory rather than a throwing out the general approach, as neither paradox was caused by incorrect statistical reasoning.

The possibility that people might have trouble reasoning about risk was however raised in work by Kahneman et al (1982) who questioned the ability of people to make accurate inferences from statistical information. This finding was observed in other pieces of

research in psychology with the added observation that people had difficulties in calculating probabilities from statistical information, particularly conditional probabilities. These results do have the potential to challenge the subjective expected utility framework since they challenge the notion that people's thinking about uncertainty obey the laws of probability. Good summaries of this topic are Gilboa (2009), Machina (1987, 2005), Starmer (2000), and Wakker (2004). Making matters even more complicated is the possibility of ambiguity about the statistical information. Imprecise knowledge of probabilities is not by itself fatal to expected utility as shown by Klibanoff et al (2005). But ambiguity combined with violations of the laws of probability surely is.

Thankfully, all is not lost. What has been found to be important for how well people calculate and reason statistically is how the statistical information is communicated to them. Simply put, people find it easier to make better inferences if information is communicated to them in certain types of forms rather than others. Framing matters as all students of behavioural economics know. In this article we present a tool that makes it easier (and thus in some sense cheaper) for people to think accurately about situations of uncertainty even in the presence of ambiguity about the information. The approach has broader applications and would be useful in the classroom as a tool to help students understand and think about probabilities, and more generally in private and public organisations by those making decisions or communicating risk information when facing uncertainty about the resulting outcomes.

2. Can People Reason Accurately in Uncertain Situations?

Given the importance of understanding how people think about risk, much research has been undertaken about how people reason statistically, particularly whether or not their reasoning is accurate, and the conditions under which their accuracy and statistical reasoning processes might be improved. As might be expected, the key issues have no simple answers, although much has been learnt along the way in trying to find the answers.

2.1 Problems People Have in Reasoning About Uncertainty

Even though correct statistical reasoning is important for people's welfare, the evidence is that people have trouble reasoning with information about risks. Research in cognitive psychology and other areas have found strong evidence of biases in statistical reasoning by people.* The human brain has evolved in a way that can give the illusion of certainty where none exists, making people prone to preferring certainty over uncertainty. Single event probabilities are prone to misinterpretation since reference classes are typically unstated, or even worse, the event is unique in which case any probability given is likely to be a guess (and not necessarily an "educated" one). People are typically confused by what conditional probabilities mean, finding it difficult to calculate and interpret them. Information given as relative risks is open to misinterpretation since it does not indicate if the numbers involved are meaningful. The *British Medical Journal* and the *Financial Times* even included columns on how poorly people seem to think about uncertain situations, particularly when having to calculate conditional probabilities.†

* A sample of this research is Birnbaum et al (1990), Chen and Crask (1998), Dougherty and Sprenger (2006), Lewis and Keren (1999), and Reyna and Brainerd (2008) .

† Watkins (2000) and the *Financial Times*, (2003), 19 June, p.21. Difficulties people have in estimating conditional probabilities has been known for over four decades , see Bauer (1972).

These findings are more than just academic curiosities. The possibility of drawing incorrect inferences from data in courtrooms, one known as the prosecutor's fallacy and the other the defendant's fallacy, was first highlighted by William Thompson and Edward Schumann in their classic 1987 article.* In the Sally Clark case an expert witness made a mistake in calculating a joint probability which was later pointed out by the Royal Statistical Society in a press release "...expressing its concern at the misuse of statistics in the courts."† The expert calculated a joint probability on the basis that the events involved were independent whereas the evidence was overwhelmingly against this assumption. The presence of systematic and predictable difficulties in reasoning with statistical information found by researchers, such as the illusion of certainty or misinterpretation of relative risks, creates strategic incentives to exploit it. For example, pharmaceutical companies have an incentive to report the use of relative risk information as it will be more likely to convince civil servants, doctors, and potential consumers that their drugs successfully treat medical conditions, increasing the demand for them and consequently their prices. This shortcoming in people's thinking also creates doubts about the subjective expected utility model to characterise people's decision making under uncertainty (see the references in Section 1). If people's understanding and calculations of probabilities violate the laws of probability then how exactly do you model their choices under uncertainty?

2.2 Natural Frequencies and the Importance of How Statistical Information is Presented and Communicated

Though the evidence strongly suggests that humans are poor at statistical reasoning all is not lost, and thankfully for economists this includes the subjective expected utility framework. Some researchers, such as Gigerenzer (2002), claim that the primary cause of the

* Thompson and Schumann (1987).

† Online. Available: www.rss.org.uk/docs/Royal%20Statistical%20Society.doc. 4 February, 2010.

miscalculation of probabilities or the misunderstanding of statistical information is because it is presented in ways that do not suit the evolutionary structure of the human brain.* In other words, as we know from behavioural economics, the framing of the information matters. Gigerenzer argues that presenting statistical information as frequencies (this is explained in Section 3.2) is far more natural for humans given their evolutionary past. Frequency based information naturally specifies a reference class and also clearly highlights that in most situations belief in certainty is an illusion. Statistical information in the form of natural frequencies suits our evolutionary past where risk information was not in normalised forms such a probabilities and percentages. Summary count or frequency based information reduces the number of mental calculations involved in working out probabilities as well as making it easier to calculate conditional probabilities. It also means people are more likely to understand what they mean in their interpersonal communication about uncertainties. Other advantages of information in a frequency form include it making clear how meaningful changes are, because the information is in absolute risk changes, and how much evidence underlies the information at hand, because it shows how many observations have occurred.

Like most things, statistical information presented as natural frequencies is not a universal panacea. For example, Chapman and Liu (2009) find evidence that a minimal level of numeracy is needed before the beneficial effects of presenting probabilistic information as frequencies occur. Barbey and Sloman (2007) argue that natural frequencies do help people to think about risk, but that how it does so is more complicated than is put forward by Gigerenzer and that it works better in some situations than others. Overall, it does seem that people are perfectly capable of statistical reasoning but the accuracy of their results is strongly dependent on how the information is presented. This feature of people's thinking is

* For other examples supporting natural frequencies as an effective way of communicating statistical information see also Brase (2008), Gigerenzer and Hoffrage (1998), Kurzenhauser and Hertwig (2006), and Sedlmeier (2002).

best captured in notion of bounded rationality in the sense of Simon (1957). The presentation of information as frequencies is relatively cheap in terms of emotional and cognitive effort for people to process, and so they calculate probabilities accurately. The presentation of information in other forms is relatively expensive to process in terms of the same effort, so calculations of probabilities are based on heuristics and other devices and as a consequence are less accurate.

2.3 Ambiguity and Imprecise Probabilities

Complicating the use of natural frequencies and other forms of communication of statistical information to improve people's thinking about uncertain situations is the presence of ambiguity. The effect of ambiguity on people's behaviour is well known in economics from the Ellsberg paradox which showed that people prefer known risks over unknown risks. This has led to a literature in economics which includes not only preferences about risk but also preferences about ambiguity. Examples of this approach are Klibanoff et al. (2005) and Baillon et al. (2010). Mukerji (2000) and especially Nau (2007) present good summaries of recent work trying to incorporate ambiguity in economics. Klibanoff et al (p. 1849) define ambiguity in way that allows them to model indifference curves as smooth and their approach is shown in this quote from their paper: "One advantage of this model is that the well-developed machinery for dealing with risk attitudes can be applied as well to ambiguity attitudes." Ambiguity in this approach essentially means people have subjective beliefs about the values of the probabilities, in effect a two-stage approach to uncertainty. There is uncertainty about states that can occur, with probabilities of these occurring, and there is uncertainty about the values of these probabilities. One result being that just as there can be risk aversion in situations of uncertainty, there can be ambiguity aversion when people are not sure about the probabilities of the risks they face.

This of course pre-supposes that what is meant by ambiguity is a lack of knowledge about the possible outcomes that can occur. But when psychologists think of ambiguity, the meaning can be wider, with people not only lacking knowledge of the values of probabilities, but also lacking knowledge about the set of possible states or not being able to calculate probabilities correctly. An example of the former are novel situations, such as new technologies. The de Havilland Comet jet crashes that occurred in 1954 is a well known case of this type.* An obvious example of the latter is the Sally Clark case. Interestingly, Mosleh and Bier (1996) show that ambiguity arising from the lack of knowledge of probabilities is consistent with the subjective theory of probability, but that ambiguity stemming from “cognitive imprecision” is not. This suggests the smooth approach to ambiguity is not necessarily a panacea in capturing ambiguity in decision making, at least not in all situations.

3. Graphical Techniques to the Rescue

What we are considering is people making decisions in uncertain situations when there exists ambiguity from cognitive imprecision plus uncertainty about the accuracy of the information being presented. We know that the framing of probabilistic information matters in how people think about it. We also know that people may have doubts about how precise the information they are getting actually is (see Fairman (2006) for examples). The evidence suggests that Gigerenzer’s frequency based presentation of data substantially reduces problems with cognitive imprecision. But the same cognitive imprecision based ambiguity

* After two years of safe operation a Comet crashed after take-off in Rome. Thirty-five people died. Flights were briefly suspended, but a second plane crashed shortly after they resumed. An intensive investigation eventually came to the conclusion that the fault lay with metal fatigue from the high speeds and high altitudes, conditions which had previously been unknown to aeronautical engineers, and which they had thus simply not considered. Stanley (1986, p. 54) commenting on this disaster reports that “At the end of the war de Havilland had been sufficiently courageous to venture into the unknown and design and build the world’s first jet aircraft. The two accidents had been the result of factors beyond the limit of contemporary knowledge...”

can make it difficult for people to cope with imprecise frequency information. Thankfully, tools do exist to help people calculate probabilities and to draw inferences from evidence correctly, that is to lower the costs of analysing information and thus increase the amount and accuracy of the analysis of it.

We know from cognitive psychology that graphical data can be superior in terms of communicating information and decision making accuracy than other forms of data presentation. For instance, Speier (2006) finds that graphic representation of data results in more accurate decisions for tasks about comparisons, trends, and such like, than tabular data (which is superior for precise numerical tasks). Graphic data in most cases resulted in faster decisions times. This finding is mirrored in Coll et al (1994). Burkell (2004) also presents evidence that graphical data (pictorial in this study) is easier to understand than numerical data. Regarding probability calculations, Cole and Davidson (1989) find that graphic representation of probabilistic information can substantially improve the time it takes to form conditional probabilities and their accuracy, more so than tabular depiction of the information. Overall, the available evidence suggests that tools that present information graphically, particularly when the decisions to be made involve deep understanding and comparative assessments and not merely exact numerical calculations, can help people make more accurate and faster decisions. Teachers of economics are well aware of the advantages of graphical over tabular methods in the presentation of the most basic Marshallian analysis of demand and supply. This advantage of graphical methods over tabular methods holds for making decisions in uncertain situations. Finally, Natter and Berry (2005) present evidence that people who actively process information are significantly more accurate in their frequency and probability estimates. This means that a graphical tool which allows people to manipulate elements of it to process and show probabilistic information to make decisions are

even more likely to make accurate decisions than if they receive the information passively.

What we do in the rest of this paper is demonstrate a graphical tool that takes advantage of these findings in cognitive psychology and elsewhere and which can be used by people to think about situations of uncertainty where ambiguity about is present. Furthermore, the tool is simple to use. Apart from making it easier for people to make better decisions, the tool would also be useful in the classroom to teach students about probabilistic information.

3.1 Statistical Background Basics

The method we use to represent Gigerenzer style natural frequency methods graphically is based on Lad's (1996) geometrical exposition of de Finetti's Fundamental Theorem of Prevision. This little known but extremely powerful theorem in statistics facilitates the identification of coherent beliefs over a “larger” finite state space that are consistent with a “smaller” number of expectations and probability assessments about operational quantities of interest related to this state space. While a general formulation of de Finetti's theorem requires knowledge of linear algebra and convexity, when the quantities are two binary valued variables, simple 2-dimensional graphical techniques suffice.

Table 1 sets out an example of the basic information we will be working with. For ease of interpretation we use a health example, but the method is applicable for any two binary variables. That is, for all situations which have an unknown state where an imperfect diagnostic test exists which gives information about the value of the state. Let S , the binary variable in the first row of the table, be the logical truth value (1 or 0) of a proposition about the health state “person A has a disease X ”. Let D , the binary variable in the second row, be

the logical truth value of a proposition about a diagnostic health test such as “person A has a positive diagnostic test for the disease X ”. We assume that truth or falsity of each proposition can be confirmed by an operational measurement, but that knowledge of the outcomes of these measurement procedures might be uncertain, in whole or in part.

Table 1: Truth Table and Natural Frequencies

	True Positive	False Negative	False Positive	True Negative
S	1	1	0	0
D	1	0	1	0
Frequency	16	4	24	56

The top two rows in the table set out the four logically possible combinations of the truth values for a pair of propositions (S,D) in the familiar, logical truth table format used in deductive logic (see Suppes (1957, p.11)). Each column of possible (S,D) values is labelled with their conventional epidemiological name: $(S,D)=(1,1)$ identifies the situation that person A has the disease and the diagnostic signal for the disease is positive (a “*true positive*”); $(S,D)=(1,0)$ identifies the situation where person A has the disease but the diagnostic signal is negative for that disease, a “*false negative*” (in this binary context “not positive” for the disease means “negative” for the disease); $(S,D)=(0,1)$ indicates that person A doesn't have the disease but the diagnostic signal for the disease is positive (a “*false positive*”); $(S,D)=(0,0)$ indicates that the person doesn't have the disease and the diagnostic signal for the disease is negative (a “*true negative*”).

The final row of the table is an illustrative set of (precise) frequency numbers of the

kind used to express uncertainties in the context of an inference task. The individual column entries for row 3 of the table are called counts or cases. An aggregate count across all logical possibilities (columns) is also pre-specified, in this case 100, but is in general a variable component of the way information about the state and the diagnostic is represented. The actual frequency numbers in Table 1 are chosen to make the initial construction of the graph representing them easy computationally. Once the method of converting from table to graph is grasped, the frequency numbers can be easily varied to suit the inference task at hand. Note that while there are four conceptually distinct, non-negative case counts, one for each column, there are really only three logically independent counts, since they must sum to a pre-specified total count. Note also that these same frequency numbers can, after simple scaling by the total count, be interpreted as probability assessments for a joint probability mass function, $P(S,D)$, on the discrete space for the two random variables, (S,D) .

3.2 Frequency Versus Probability Representations of Probabilistic Information

The truth table just presented is simply a vehicle for representing relevant precise numerical information and it is itself neutral between frequency or probability formats. It is thus unable to help answer the question of whether ordinary intelligent people make better inferences using a frequency format or a probability format; “format” here meaning “representation”. But before proceeding any further it makes sense to first explain the difference between a frequency format and a probability format. Gigerenzer and Hoffrage (1995) explain the difference between a *standard probability format* and *standard frequency format*, in the context of an inference task specified in English textual terms as:^{*}

^{*} The sentences are quotes from Table 1 of Gigerenzer and Hoffrage (1995) with the italicized words and frequencies indicating changes from the original in order to conform to the health example and frequency numbers in our example. For example, our “disease X” replaces their “breast cancer”, and so on.

Standard probability format

The probability of *disease X* is 20 percent for women at age forty who participate in routine screening. If a woman has *disease X*, the probability is 80 percent that she will get a positive *diagnostic test*. If a woman does not have *disease X*, the probability is 30 percent that she will also get a positive *diagnostic test*. A woman in this age group had a positive *diagnostic test* in a routine screening. What is the probability that she actually has *disease X*? ___%

Standard frequency format

20 out of every 100 women at age forty who participate in routine screening have *disease X*. 16 of the 20 women with *disease X* will get a positive *diagnostic test*. 24 out of the 80 women without *disease X* will also get a positive *diagnostic test*. Here is a new representative sample of women at age forty who got a positive *diagnostic test* in routine screening. How many of these women do you expect to actually have *disease X*?

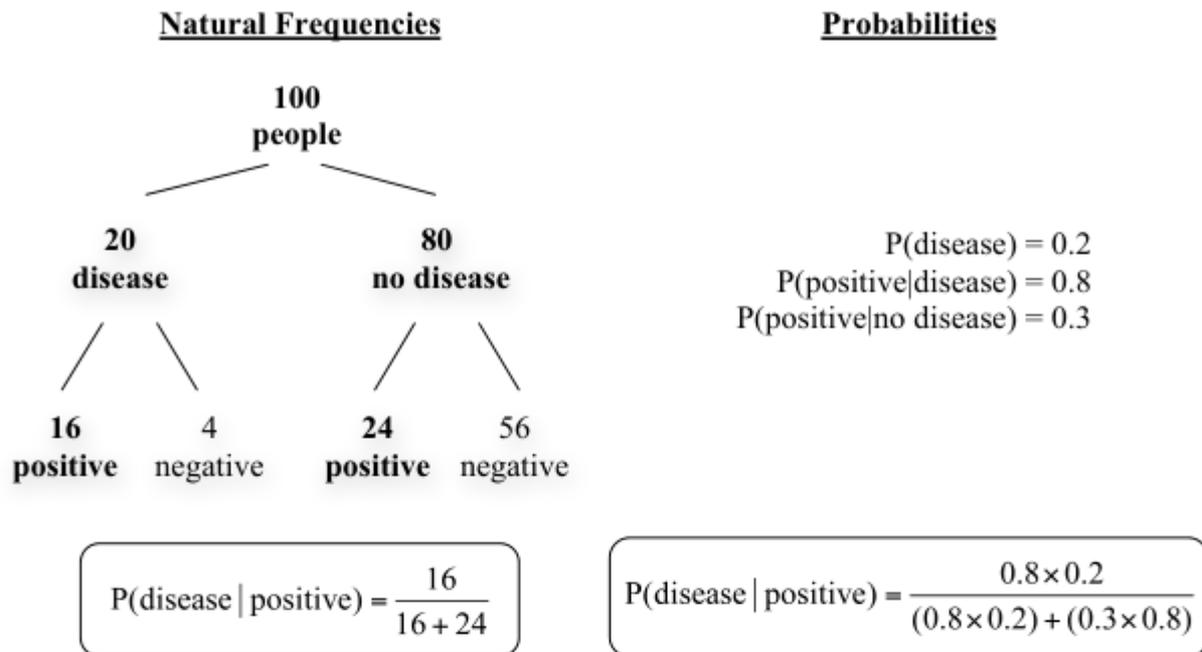
Posing an inference problem in text form is not the only way quantitative information in frequency or probability formats can be, or is, presented to subjects in experimental research. Figure 1 below, based on Figure 4-2 in Gigerenzer (2002, p.45), but adapted to our tabular frequencies, illustrates the difference between the two formats in a tree diagram superimposed on tabular information (the bold emphasis, serving to draw attention to particular numbers, is in the original). The upper panels in Figure 1 for each format are different, albeit logically equivalent, ways of representing information relevant to the inference task at hand: what are the chances of having a disease given a positive test result,

on the basis of the precise information provided? The boxes in the lower panels show the kinds of calculations that need to be made to arrive at a correct inference (40 percent is the posterior probability of disease given a positive diagnostic test result). Once someone learns to “read” the table and the superimposed tree in the frequency format in the left-hand panel, so that the $16+24=40$ cases or counts of positive diagnostics can be readily identified and classified into those diseased or not, then the task of working out what fraction, proportion, or number of these 40 actually having the disease is greatly simplified. Certainly in a comparative sense this computational task is much easier than the inverse probability calculations required in the right hand probability format panel. As Gigerenzer (2002, p.45) puts it: “The representation does part of the computation.” The dynamic graphical method for representing uncertainties we present later on in the paper will take this insight farther where the representation itself does the whole computation.

3.3 Graphically Representing Natural Frequencies

As noted previously, once the total count is specified there are really only three logically independent entries in the column entries of Table 1 available for representing uncertainty about (S,D) . But three distinct cell counts aren't the only, nor the most useful, way of identifying the three relevant bits of information. For example the text formats in Section 3.2 specify three other numbers as a way of communicating information relative to reference: a **sensitivity number**, a **specificity number**, and a **base rate number**. These three numbers mix and match uncertainties both about the state variable S and about the diagnostic variable D viewed conditional on information about the state variable S . The sensitivity number expresses uncertainty about whether the diagnostic test will be positive, or $D=1$, assuming that $S=1$ is true. In our example this is the 80 percent probability asserted in the sentence “If a woman has *disease X*, the probability is 80 percent that she will get a positive *diagnostic*

Figure 1: Frequency and Probability Representation Formats



test.” The specificity number expresses an uncertainty about whether the diagnostic test will be negative, or $D=0$, assuming that $S=0$ is true. In our example this is a 70 percent probability, or 1 minus the 30 percent probability asserted in the sentence: “If a woman does not have *disease X*, the probability is 30 percent that she will get a positive *diagnostic test*.” Probabilistically sophisticated people understand these numbers as conditional probabilities, $P(D=1|S=1)$ for sensitivity and $P(D=0|S=0)$ for specificity. The base rate number for the state variable characterizes uncertainty about the binary state variable S in the absence of, or prior to, learning any diagnostic information D . In our example this is the 20 percent probability asserted in the sentence “The probability of *disease X* is 20 percent for women at age forty who participate in routine screening.” Probabilistically sophisticated people understand the

base rate number for the state variable as a marginal or unconditional probability, $P(S)$.^{*} Probabilistically sophisticated people also understand that specifying a full joint probability distribution, $P(S,D)$, as in the cells of Table 1, or specifying the sensitivity number $P(D=1|S=1)$, the specificity number $P(D=0|S=0)$, and the base rate number $P(S=1)$, are logically equivalent ways of saying the same thing. But, what about probabilistically unsophisticated people? Or, probabilistically sophisticated people in situations where there is ambiguity about the source of the probabilistic information? What we do next is present a graphical method to make it easier for such people who find themselves in these situations to perform statistical calculations, particular statistical calculations. In essence the method lowers the costs of performing the underlying statistical calculations, meaning in situations of bounded rationality there is less reliance on previously less costly but less accurate heuristics, with commensurate increases in the speed and accuracy of the calculations.

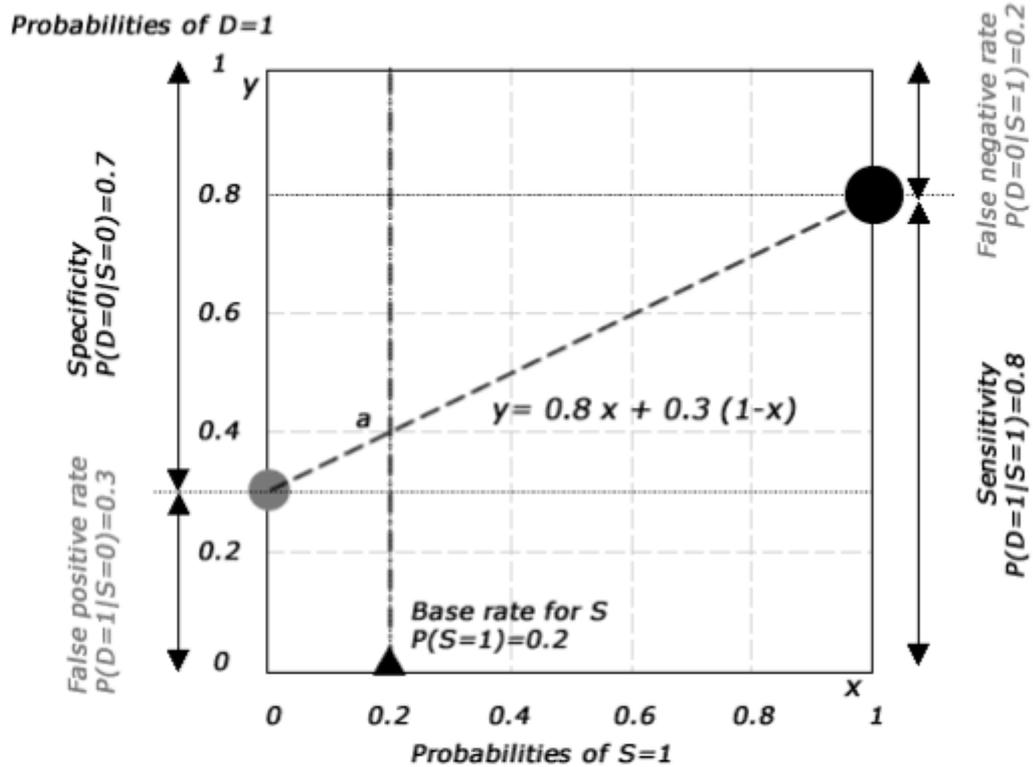
Figure 2 below is the first step of a three stage process in how to represent a specific table of natural frequencies graphically. Conditional or marginal probabilities (or counts) for the presence of the disease, $S=1$, are plotted in the x-axis direction and conditional or marginal probabilities (or counts) for the diagnostic D being positive, $D=1$, are plotted in the y-axis direction. At first we will view the graphical representation of the frequency table simply as a collection of three dots, but a second more insightful approach is to connect the dots and view the frequency table as the intersection of several linear coherency constraints on beliefs.

^{*} The many nuances to the concept of “probabilistic sophistication” are well discussed in Gilboa (2009). In our paper we use the term simply to describe an ability to quantify uncertainties and to reason consistently within the boundaries of the laws of probability.

Figure 2: Graphical Representation of Sensitivity, Specificity, and Base Rate
Derived from a Natural Frequency Table

Truth Table and Natural Frequencies

	True Positive	False Negative	False Positive	True Negative	
S	16	4	0	0	
D	16	0	4	0	
Frequency	16	4	24	56	100



The first two columns of the table in Figure 2, where $S=1$, show that $16+4=20$ out of 100 cases are associated with the proposition being true that person A has the disease. This, gives a base rate, $P(S=1)$, equal to 20 out of 100 or 20 percent, shown as a triangle on the x-axis along a vertical line (dash-dot) at the point (0.2,0) on the horizontal axes in Figure 2. Still focusing on the first two columns and the 20 cases where the disease is present, the sensitivity of the diagnostic test, $P(D=1|S=1)$, is 16 out of 20 or 80 percent, shown as a circle

on the right hand margin of the graph with coordinates (1,0.8). The diagnostic test is good, albeit imperfect, at detecting the presence of the disease state. If the disease is there with 100 percent surety ($P(S=1)=1$ along the right hand margin), the diagnostic picks this up with a high (80 percent) chance in this circumstance. But the diagnostic test does get it wrong with a 20 percent chance (a false negative type of wrong). This false negative rate, $P(D=0|S=1)=0.2$, is also indicated by the circle on the right hand margin of the graph, reading down from the top right corner. That *one* dot indicates *two* interesting uncertainties about the diagnostic D , under the assumption that the disease is there with 100 percent surety ($P(S=1)=1$ along the right hand margin).

Using the last two columns of the table where the disease is absent, or $S=0$, 24 cases are false positives and 56 are true negatives. These two numbers determine the false positive rate and the specificity of the test. The specificity of the test, $P(D=0|S=0)$, equal to 56 out of 80, or 70 percent, is shown as a circle on the left hand margin of the graph at a height of 30 percent, the false positive rate of the test. The diagnostic test is good, albeit imperfect, at detecting an absence of the disease state. If the disease is absent with 100 percent surety ($P(S=1)=0$ along the left hand margin), it picks this up with a high (70 percent) chance. But the diagnostic test does get it wrong in these circumstances with a 30 percent chance (a false positive type of wrong). The false positive rate, $P(D=1|S=0)=0.3$, is indicated by the circle on the left hand margin of the graph, reading up from the origin. The specificity of the diagnostic test is indicated by the same circle on the left hand margin of the graph, reading down from the top left corner. That *one* dot indicates *two* interesting uncertainties about the diagnostic D , under the assumption that the disease is absent with 100 percent surety ($P(S=1)=0$ along the left hand margin).

The dashed line joining the two circular dots on the left and right hand margins, with equation $y = 0.8x + 0.3(1-x)$, is a linear coherency constraint on overall probability assessments.^{*,†} It is based on the mathematical fact that the base rate for the diagnostic test, the unconditional probability of having a positive diagnostic, $P(D=1)$, should be a weighted average of the chances of having a positive diagnostic when disease is present, $P(D=1|S=1)$, and the chances of having a positive diagnostic when the disease is absent, $P(D=1|S=0)$, with weights being the chances of presence, $P(S)$, or absence, $1-P(S)$, of the disease state. In algebraic terms:

$$P(D = 1) = P(D = 1|S = 1)P(S = 1) + P(D = 1|S = 0)(1 - P(S = 1)) .$$

Letting $P(D=1)$ be the y-axis variable and $P(S=1)$ be the x-axis variable, and using $P(D=1|S=1)=0.8$ and $P(D=1|S=0)=0.3$ from the table, the line $y = 0.8x + 0.3(1-x)$ traces out all combinations of (x,y) . In this context it shows all combinations of $(P(S=1)$ and $P(D=1))$ consistent with the given sensitivity and specificity numbers of 0.8 and 0.7. For example when the base rate for the state variable is $P(S=1)=0.2$, the base rate for the diagnostic will be $0.8*0.2 + 0.3(1-0.2) = 0.4$, the intersection of the dashed line and the vertical dot-dashed line through the base rate triangle on the graph. The important point to realize is that once the two conditional probabilities, the sensitivity and the specificity, are fixed, the permissible combinations of $P(S)$ and $P(D)$ must lie along that dashed line.

Exactly the same logic can be applied to posterior inferences about the state variable

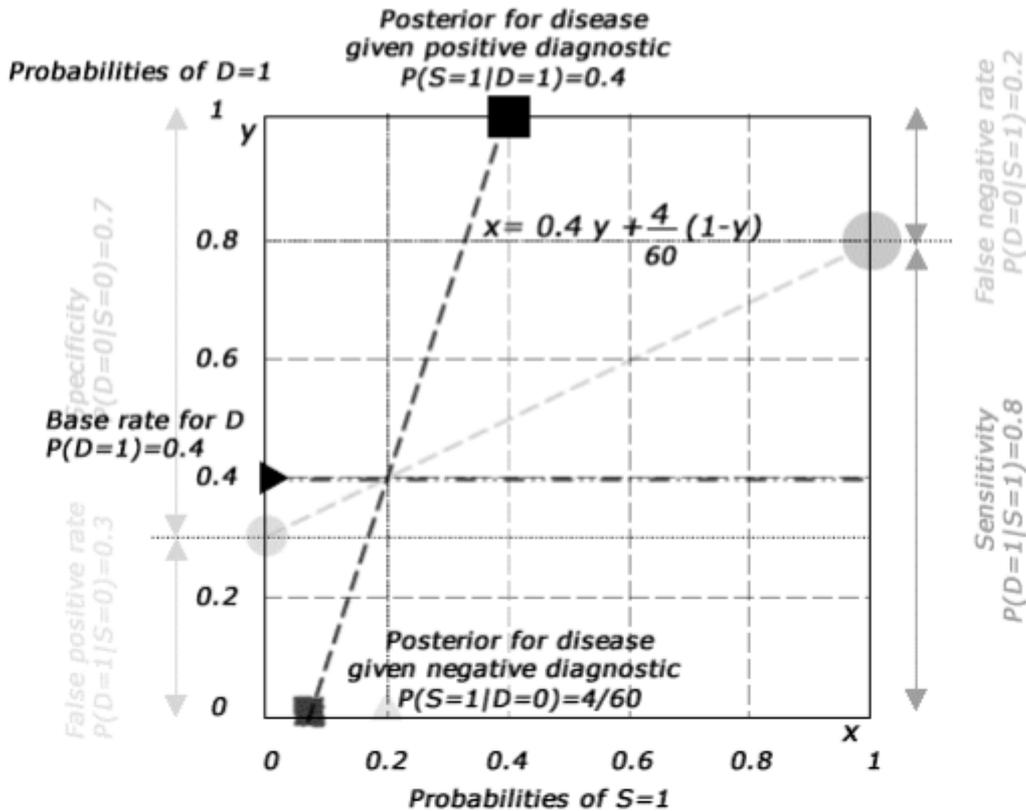
^{*} Coherency here can be interpreted simply in the formalist sense of consistency with the laws of probability or in the operational subjective sense of being unwilling to assert probabilities that would make you a sure loser. See Ladd (1996) for more about this concept.

[†] A short webcast explainaing how to download and use the dynamic interface can be found here: <http://uctv.canterbury.ac.nz/post/4/1049>. As explained in the webcast and associated documentation, the Mathematica Player used to run the interactive demonstration is freely downloadable and no prior experience with Mathematica is required.

Figure 3: Graphical Representation of Posterior Inferences Based on the Natural Frequency Table

Truth Table and Natural Frequencies

	True Positive	False Negative	False Positive	True Negative	
S	1	1	0	0	
D	1	0	1	0	
Frequency	16	4	24	56	100



being true, $S=1$, given diagnostic test information, either positive, $D=1$, or negative, $D=0$.^{*} The squared box in Figure 3 along the *upper* boundary x -axis at 0.4 is the posterior or inverse inference about the chances of the disease state being present when a positive diagnostic is observed, $P(S=1|D=1)$.[†] It is derived from columns 1 and 3 in the table in Figure 3, where $D=1$, with 16 out of the total $16+24=40$ cases of a positive diagnostic being associated with

^{*} By “posterior” we mean after learning the outcome of the diagnostic test but while still being uncertain about the state variable S .

[†] $P(S=1|D=1)$ is known as an inverse conditional probability since the role of S and D and “inverted” compared to the sensitivity assessment $P(D=1|S=1)$.

the disease being present and the other 24 cases associated with the disease being absent in these circumstances. The smaller squared box in Figure 3 along the *lower* boundary x-axis at $4/60 \approx 0.067$ is the posterior or inverse inference about the chances of the disease state being present when a negative diagnostic is observed, that is $P(S=1|D=0)$. It derives from columns 2 and 4 in the table, where $D=0$, with four out of the total $4+56=60$ cases of a negative diagnostic being associated with the disease being present and the other 56 cases associated with the disease being absent in these circumstances. The base rate for the diagnostic test being positive, $P(D=1)$, is plotted as a triangle and horizontal dash-dotted line at height 0.4. The line $x = 0.4y + 4/60(1-y)$ between the two conditional posterior inferences (the dot-dashed line between the squared boxes) is another linear constraint on marginal probabilities $(P(S=1), P(D=1))$. It derives from the mathematical fact that the marginal or unconditional probability for the disease being present, the base rate for the disease, or $P(S=1)$, must be an appropriate weighted average of the two conditional probabilities of having the disease, $P(S=1|D=1)$, for those with positive diagnostic test results, and $P(S=1|D=0)$ for those without positive diagnostic tests. The base rate for the diagnostic test (distinguished from the base rate for the state variable S) being positive, or $P(D=1)$, determines the appropriate weight to be used. In algebraic terms:

$$P(S = 1) = P(S = 1 | D = 1)P(D = 1) + P(S = 1 | D = 0)(1 - P(D = 1)).$$

Letting $P(D=1)$ be the y-axis variable and $P(S=1)$ be the x-axis variable, and using $P(S=1|D=1)=0.4$ and $P(S=1|D=0)=4/60$ from the table, the line $x = 0.4y + 4/60(1-y)$ traces out all combinations of (x,y) . That is it traces out all combinations of $(P(S=1)$ and $P(D=1))$ consistent with the given posterior probabilities of $P(S=1|D=1)=0.4$ and $P(S=1|D=0) = 4/60$ from the frequency table.

Figure 4: Graphical Representation of All Interesting Inferences Based on the Natural Frequency Table

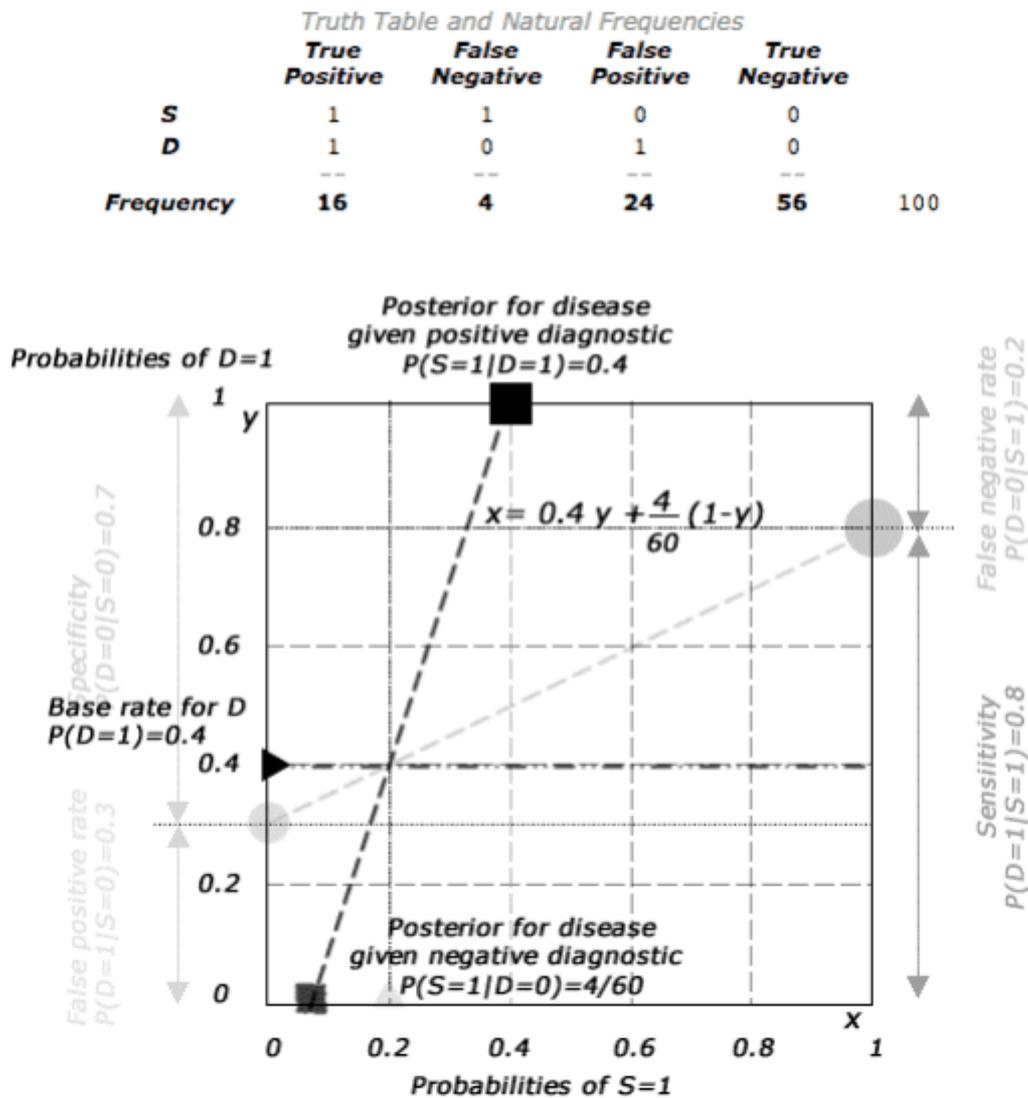


Figure 4 combines all of the information in Figures 2 and 3 into one, albeit complicated, picture that expresses many of the relevant uncertainties explicit and implicit in Table 1: base rates for diseases and diagnostics, sensitivities, specificities, false negative and false positive rates, as well as posterior inferences for diseases given positive or negative bits of diagnostic information. Not all of these uncertainties are necessarily of interest. For example the questions posed above by Gigerenzer focus on inferring one posterior probability only, that being the chances of having a disease conditional on receiving a positive diagnostic

test result. In Section 4 we will show a pen and pencil back of the envelope way to graphically compute both an answer to this question and also to investigate how the answer changes in response to variations in the inputs used to construct it (the base rate, the sensitivity, and the specificity). But with modern interactive dynamic methods these tasks can be made very simple in user friendly representations. If a Gigerenzer style natural frequency table does part of the calculation for a subject on the fly, then a dynamic graphical representation removes all of the calculation effort.

The sequence of four graphs in Figures 5a through 5d are static screen shots from the dynamic *Mathematica* interface we have created. Figure 5a essentially reproduces Figure 4 above, except we have scaled the total counts to sum to 1,000 rather than 100. The sliders on the left-hand side control the base rate, the sensitivity, and the specificity of the test. The top set of sliders can be manipulated one at a time or all at once, and the frequency table and corresponding graphic will change. The bottom set of sliders set a benchmark level for the base rate, the sensitivity, and the specificity of the test on the graphic, so that before-after comparisons can be easily and visually made. The posterior inferences about the health state variable S conditional on either positive or negative diagnostic information are shown automatically in the graph by the squared boxes. This particular interface is designed to focus attention on the posterior probability of a disease given a positive diagnostic — the large blue square on the top x-axis boundary labelled $P(S|D=1)$. Note that to keep the notation simple we are using the symbol S for both the binary state indicator variable and the proposition “ $S=1$ ” writing $P(S|D=1)$ for $P(S=1|D=1)$.

Figure 5a: Interactive Mathematica Interface — at Default Values for the Table and the Graph

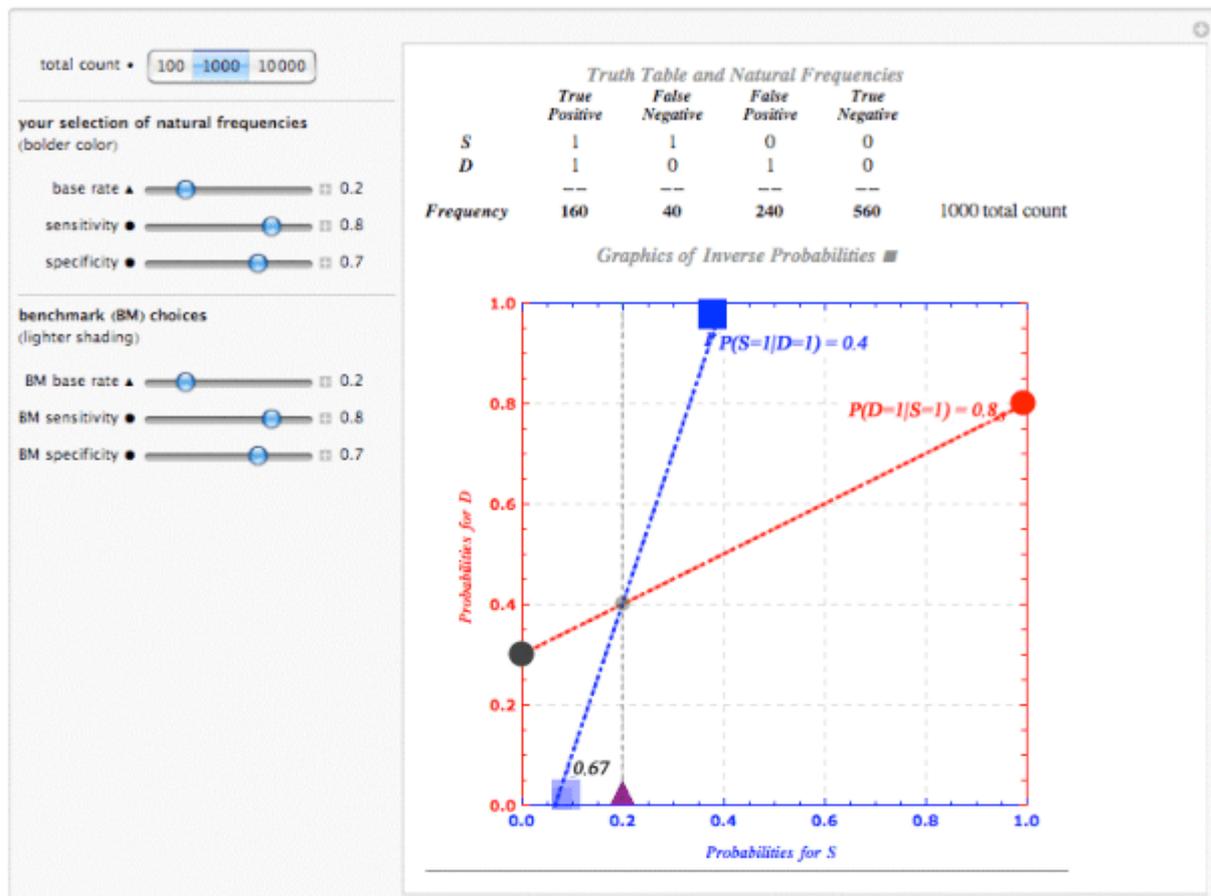
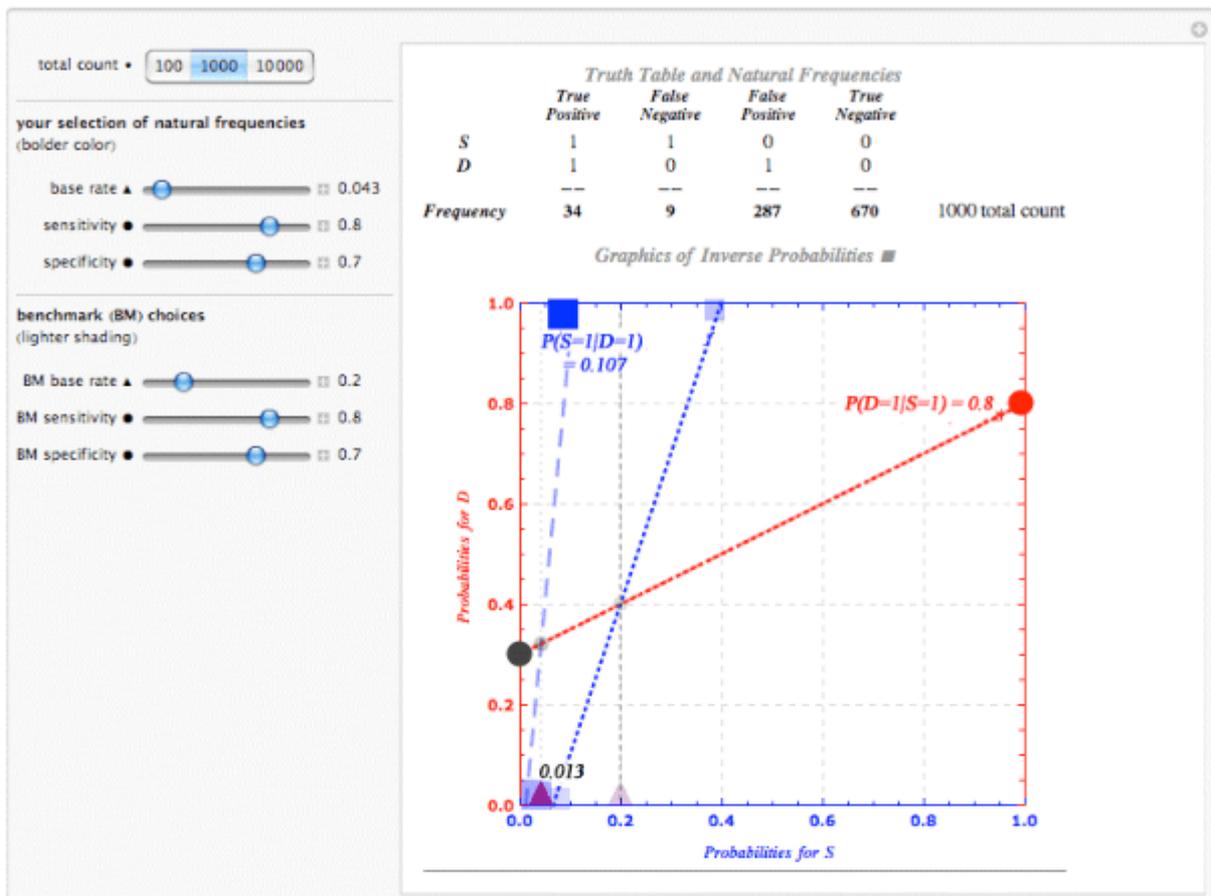


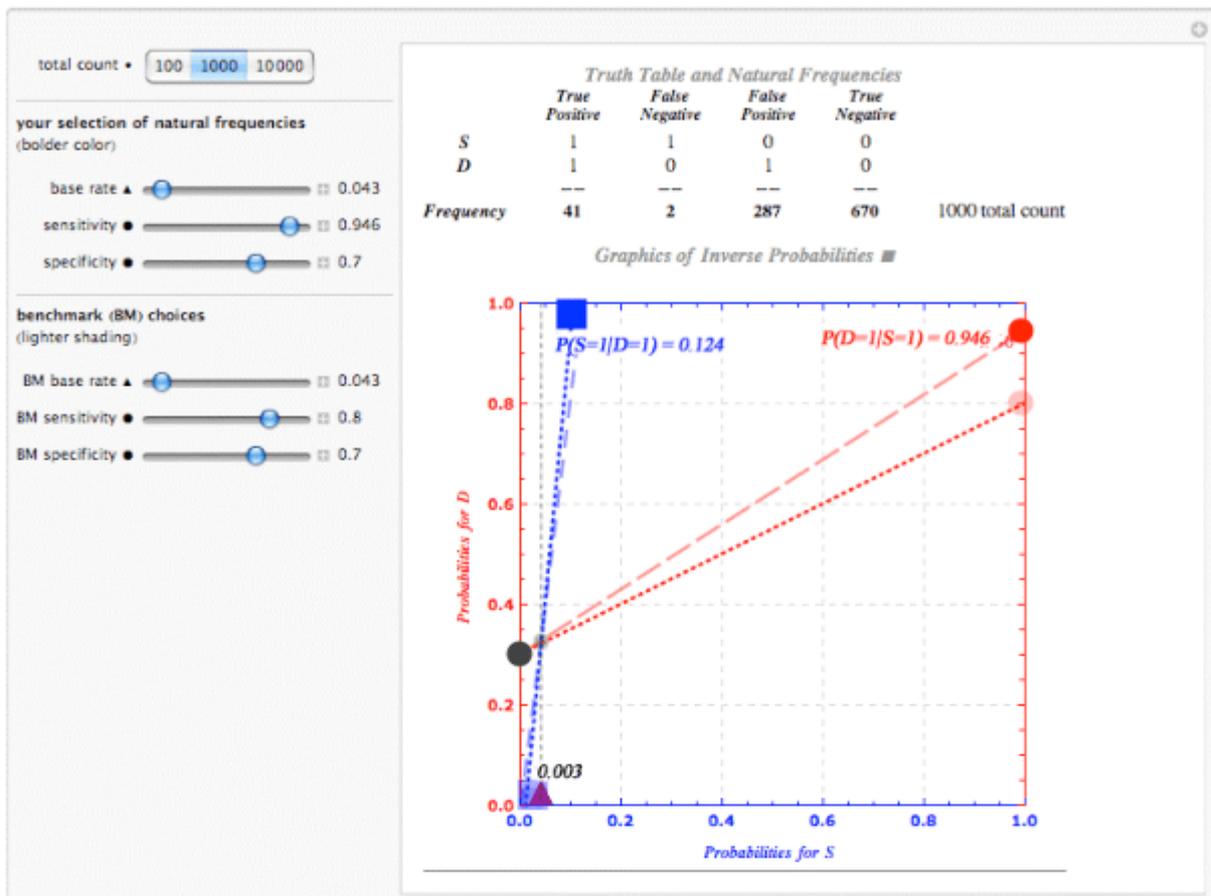
Figure 5b provides one example of how to use the interactive interface — a reduction in the base rate of the disease from 20 percent to 4.3 percent. Both the sensitivity and the specificity of the test remain unchanged. The table changes to reflect the new lower base rate, and the posterior inference drops dramatically from its previous value of 40 percent down to just under 11 percent. The graphic clearly shows this quantitative change by keeping dotted lines for the original posterior coherency constraint and the smaller squares as the benchmark posterior inference (from a base rate of 20 percent, a sensitivity of 80 percent, and a specificity of 70 percent), then changing to the dashed lines for the new posterior coherency constraint and the larger squares at the new posterior inference (from a base rate of 4.3 percent, a sensitivity of 80 percent, and a specificity of 70 percent). Intuitively, for a given

Figure 5b: Interactive Mathematica Interface Showing the Effect of Reducing the Base Rate for Fixed Sensitivity and Specificity



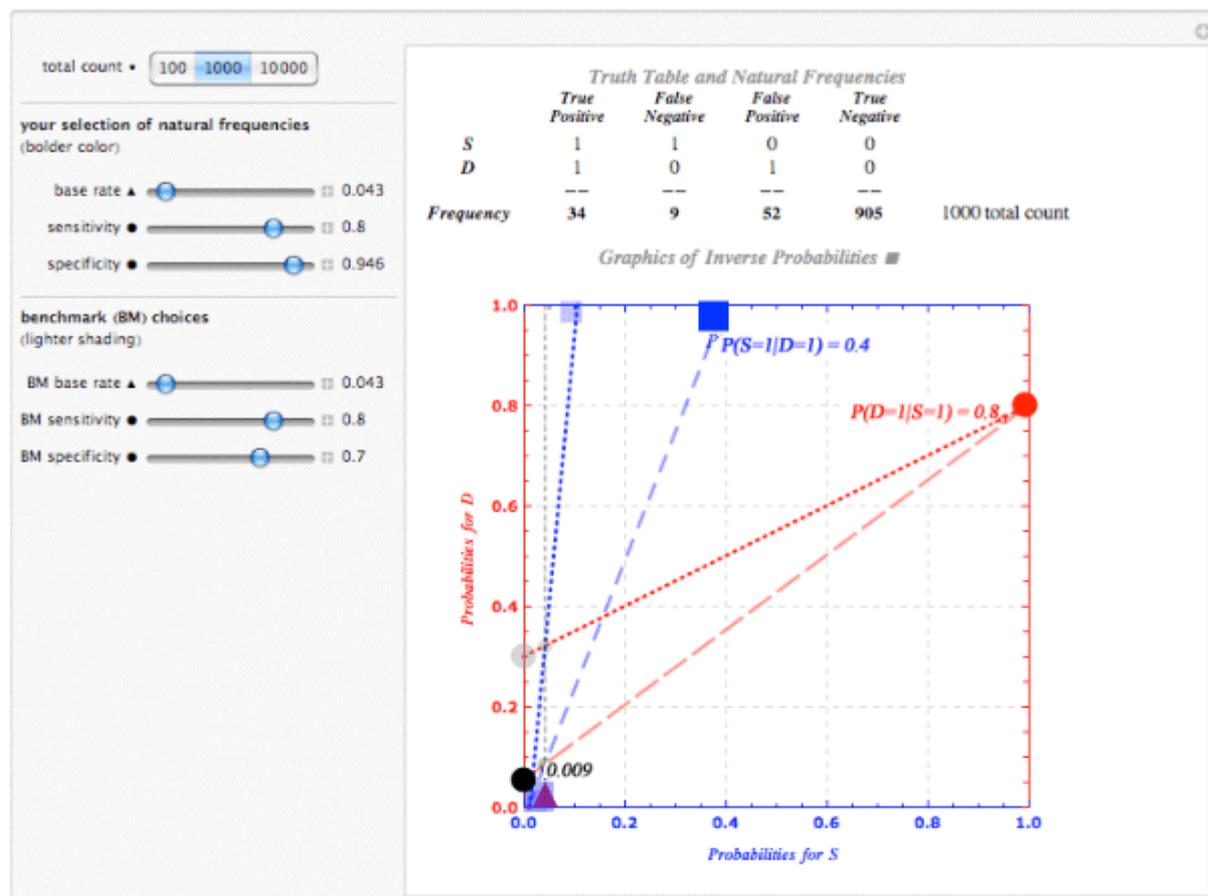
imperfect test, variations in the base rate of the disease from 1 in 5 down to about 1 in 25 can bring about a dramatic reduction in the posterior probability of having the disease given a positive diagnostic test result. And the natural frequency representation in the table makes it clear as to why: the vast majority of the positive signals, columns 1 and columns 3, come from false positives (287) rather than from true positives (34). Further experimentation in the dynamic interface with changing the base rate, and only the base rate, quickly shows how important the level of the base rate is to posterior inferences for a test of precisely known specificity and sensitivity.

**Figure 5c: Interactive Mathematica Interface Showing the Effect on Posterior Inference
Given a Positive Diagnostic from a Test with Improved Sensitivity**



But what if either the sensitivity or specificity, or both, are also ambiguous? Figure 5c shows the effect on an initial posterior inference, $P(S=1|D=1)$, of 10.7 percent, of improving the sensitivity of the test $P(D=1|S=1)$ from 80 percent to just under 95 percent (starting from a reference base rate of 4.3 percent, a sensitivity of 80 percent, and a specificity of 70 percent). There is only a modest increase in the posterior inference $P(S=1|D=1)$ from just under 11 percent to just over 12 percent. Again, the frequency table provides an insight into why this occurs. It shows that the improved sensitivity has done nothing to change the large number (287) of false positives in column 3 (compare column 3 in Figure 5b with column 3 in Figure 5c), although it has slightly increased the true positives to 41 (compare column 1 in Figure 5b with column 1 in Figure 5c).

**Figure 5d: Interactive Mathematica Interface Showing the Effect on Posterior Inference
Given a Positive Diagnostic from a Test with Improved Specificity**



What if the specificity of the test is improved? Figure 5d shows the effect on an initial posterior inference, $P(S=1|D=1)$, of 10.7 percent, of improving the specificity of the test $P(D=1|S=1)$ from 70 percent to just under 95 percent (starting from a reference base rate of 4.3 percent, a sensitivity of 80 percent, and a specificity of 70 percent). There is a dramatic increase in the posterior inference $P(S=1|D=1)$ from just under 11 percent to 40 percent. Again, the frequency table provides insight into why the improved specificity dramatically reduces the number of false positives in column 3, (compare column 3 in Figure 5b with column 3 in Figure 5d).

Many other interesting parameter changes, one at a time or jointly, can be made and

easily compared within the dynamic interface, with all of the computations suppressed and without the cognitive burden of processing arrays of tabular information.

4. A “Back of the Envelope” Approach

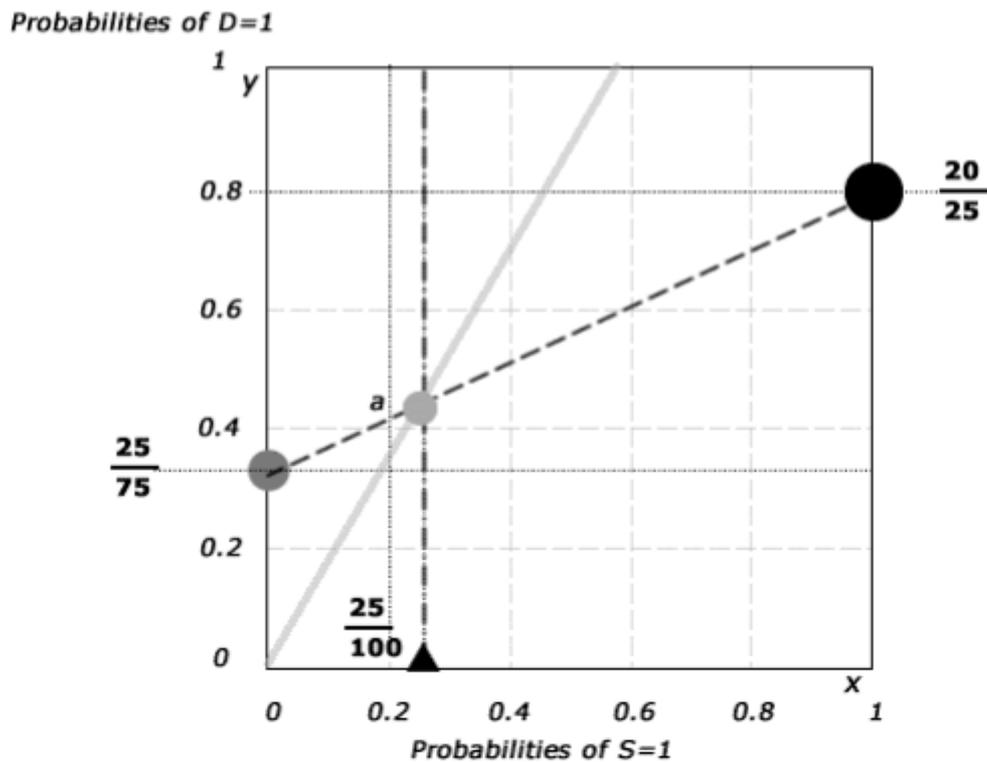
What if a dynamic Mathematica interactive interface isn't available to help in a posterior inference task? It turns out that a pencil and paper approach can still be very useful, both for representing an initial “ballpark” but precise assessment of uncertainty between signal and diagnostic and for exploring ambiguities — a lack of precision — in the three key components of that uncertainty: in any of the sensitivity and the specificity of the diagnostic testing system, or the base rate for the underlying state. Figure 6 illustrates the first step in the graphical procedure, a case of simply connecting the dots! It starts with a “simple” frequency table, where every number is a multiple of 5, and the total is 100. Think of this as assessing uncertainties in broad ranges of 5 percent. The sensitivity of the diagnostic test (concentrate on the relative size of columns 1 and 2) is assessed here at $20/(20+5)$ or 80 percent, plotted on the right hand margin as a large bullet (on the right hand margin to correspond with the idea of being 100 percent sure that the state variable $S=1$). The specificity of the diagnostic test (concentrate on the relative size of columns 3 and 4) is assessed here at $50/(25+50)$ or approximately 67 percent, plotted on the left hand margin as a slightly smaller black bullet (on the left hand margin to correspond with the idea of being 100 percent sure that the state variable $S=0$). The base rate for the state variable is assessed at $(20+5)/100$ or 25 percent, and plotted as a triangle on the x-axis at 25 percent. The first coherency line is then drawn between the sensitivity and the specificity points, and a vertical line is drawn through the base rate point, effectively re-creating Figure 2 in Section 3, for a different set of frequency counts. In keeping with a “quick-and-dirty” approach common with pencil and paper

Figure 6: The First Step in a Pencil and Paper Approach to Finding the Posterior

Inference $P(S=1|D=1)$

Truth Table and Natural Frequencies

	<i>True Positive</i>	<i>False Negative</i>	<i>False Positive</i>	<i>True Negative</i>	
<i>S</i>	1	1	0	0	
<i>D</i>	1	0	1	0	
<i>Frequency</i>	20	5	25	50	100



techniques we have not reduced the fractions on the graph nor labelled each of the points in the language of probability notation, as we did in Figure 2.

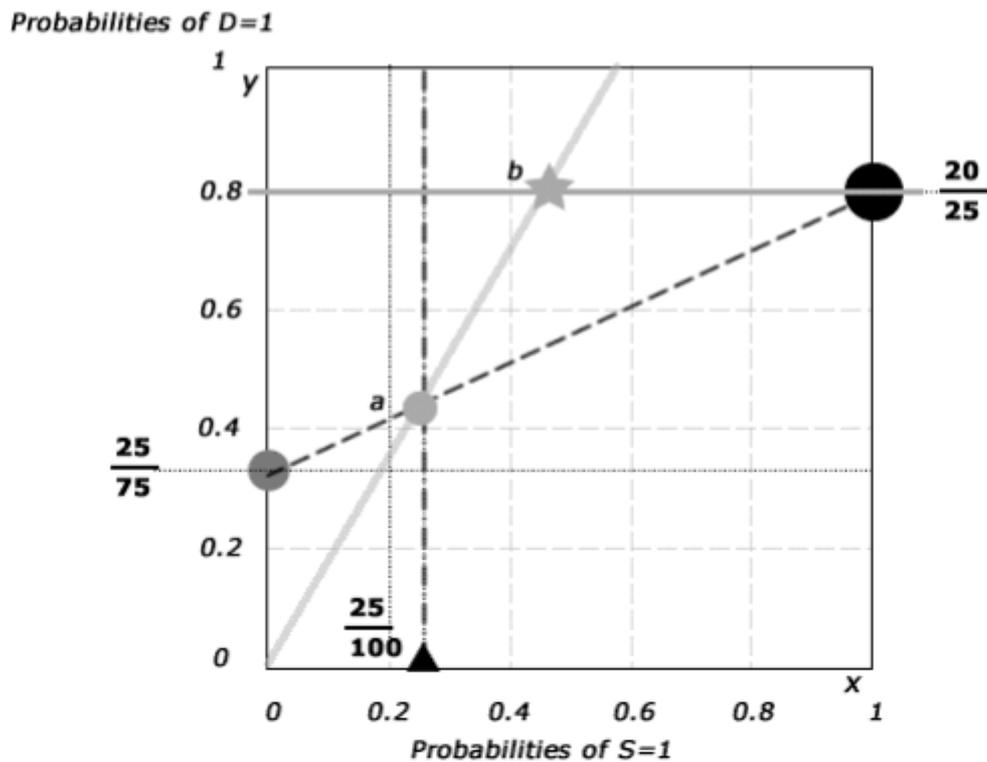
There is an additional line, shown in gray, that has been added to the graph. This ray through the origin to point “a”, the point identifying the coherent base rates for the state and the diagnostic, $(P(S=1), P(D=1))$, should be thought of for now as simply an artificial, mechanical reasoning aid used to make posterior inferences. Figure 7 provides the second

Figure 7: The Second Step in a Pencil and Paper Approach to Finding the Posterior

Inference $P(S=1|D=1)$

Truth Table and Natural Frequencies

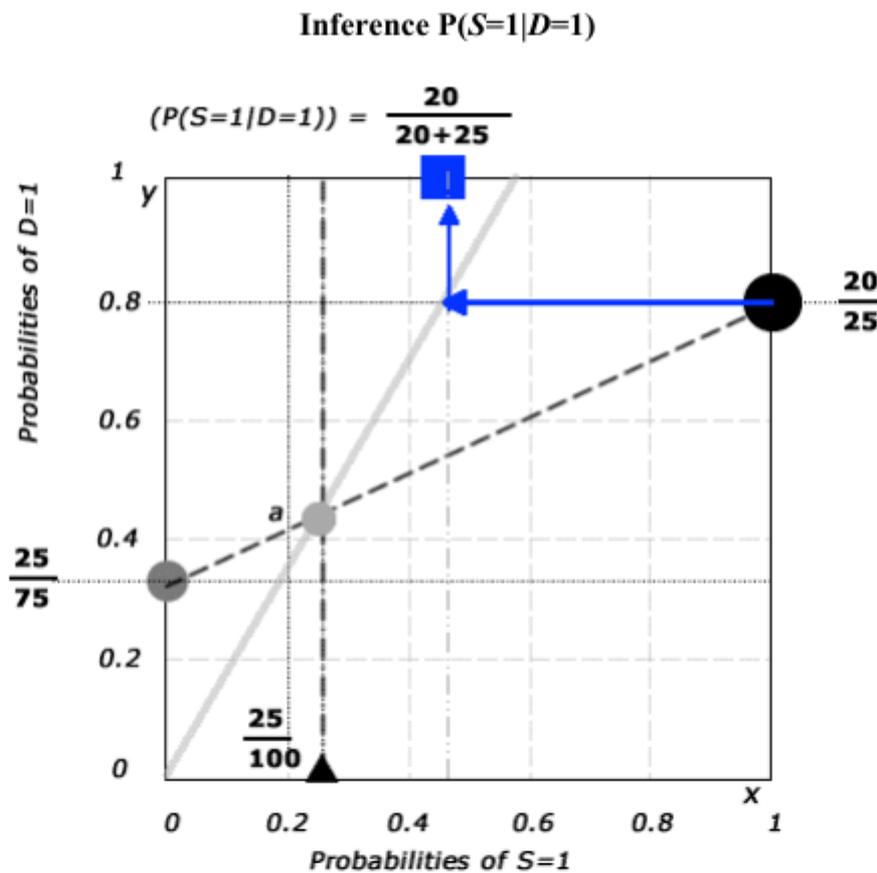
	True Positive	False Negative	False Positive	True Negative	
S	1	1	0	0	
D	1	0	1	0	
Frequency	20	5	25	50	100



step in the pencil and paper approach. Plot a horizontal line at the height of the specificity point, find the intersection with the artificial drawing aid line, the gray ray through the origin and through “a”. In Figure 7 that point is marked with a star and labelled “b”.

Figure 8 augments the sketch to date with critical information from the table of frequencies. It puts a vertical line through point “b” and identifies the intercept of that line with the top x-axis with a square box. The position of that square box is the posterior

Figure 8: The Third Step in a Pencil and Paper Approach to Finding the Posterior



probability we are looking for, $P(S=1|D=1)$. The number plotted there is readily calculated from the frequency table by focusing on the relative sizes of the frequencies in columns 1 and 3 (where $D=1$). There are a total of $20+25$ cases in those two columns, and $20/(20+25)$, slightly less than half, or approximately 44 percent using a calculator, have the disease ($S=1$). The remaining 25 are false positive signals ($S=0$). The arrows in Figure 8 are a shortcut way of finding that square box. Starting from the sensitivity point, or the large black bullet on the right margin, move backwards horizontally until hitting the artificial ray, then move vertically upward to the top x-axis

Why does this work? Well, algebraically, a modification of Bayes theorem shows that:

$$\frac{P(D = 1)}{P(S = 1)} = \frac{P(S = 1 | D = 1)}{P(D = 1 | S = 1)}$$

Along the ray through the origin and “a”, that left hand side ratio $P(D=1)/P(S=1)$ is a constant, say k . Let x stand for the probability we are interested in finding, $P(S=1|D=1)$, and y stand for the sensitivity, $P(D=1|S=1)$. Then we have the equation $k = x/y$. Given y at the level of the specificity, we solve for x using this equation. Why do we plot this number on the top x-axis? Along that line it is for sure that $P(D=1)$ which is the conditioning statement in the posterior inference $P(S=1|D=1)$.

Figure 9 shows how this pen and pencil technique can be harnessed to investigate ambiguity in the knowledge of the sensitivity and the specificity of a diagnostic test, and of the base rate of the disease, either one element at a time, or altogether. Figure 9 grays out the graphical representation of the original frequency specification; the one used to identify, precisely, the sensitivity, the specificity, and the base rate of Figure 8. Suppose that there is ambiguity about the specificity number. Perhaps the diagnostic test method is not as good as recognizing that in two out of three cases (or in this case 50 out of 75) where there is no disease, $D=0$, but has a specificity of only 50/50, as good (or as bad) as a coin toss. The table introduces some new frequencies to reflect this — 37.5 or half of the original 75 counts of $S=0$ are shared equally between $D=1$ and $D=0$ instead of in the ratio 25 to 50 — but keeps the base rate and the sensitivity of the test the same (nothing changes in columns 1 and 2).

The new graphic that represents this frequency information is constructed as an overlay. The first new point to construct is the new sensitivity, the bolder circle on the left hand margin, then the dashed coherency line between the sensitivity and specificity numbers, then the new artificial aid, a ray through the origin and point “A” where the base rate line and

much (from 44 percent to 35 percent). Again, the frequency table provides insight into why — the change in the specificity here increased the number of false positives without affecting the number of true positives, and so decreased the chance that a positive diagnostic test result truly indicates a diseased state. While we have only indicated the effect (on posterior chances of having the disease when a positive diagnostic test is observed) of a drop in specificity, ambiguity might also prevail on the up-side — the test might have a specificity better than its original 70 percent. Now, a similar diagram will reveal that the posterior probability has increased. Not knowing the specificity precisely, but within bounds, it is possible with reasonable cognitive effort to figure out the associated bounds on the posterior chances of having the disease. One ambiguity (in specificity) turns into another ambiguity (in posterior inferences), but the decision maker is not living under the illusion of certainty that a single precise frequency table would provide.

5. Summary

Every day, people have to make important decisions in risky situations. These can include whether to undertake costly medical treatments for potential diseases, or to incur the costs of shutting down production lines for maintenance or replacement of machinery to avoid product quality defects, or to incur costs from shutting down a nuclear reactor to prevent excessive strain on the pressure vessel or in the worst case the uncontrolled release of radioactive material. In all these situations, diagnostic tests or their equivalent are available about the state variables of interest, however imperfect the information from the tests may be. Making decisions based on these tests is problematic when information underlying inferences about the value of the state variables is ambiguous (or imprecise). Gigerenzer's natural frequency methods are undoubtedly a step in the right direction in trying to improve the

ability of practical people to draw sound inferences from diagnostic information. But being the exact numbers they are, they also (falsely) offer an illusion of certainty (about the frequency numbers themselves) in their trying to convey uncertainty about an underlying state variable of interest

This paper presents a software tool that can aid in the making of decisions in these types of situations. It does this by allowing people at a low cost to incorporate ambiguity they might have about the probabilistic information they are using to make inferences about the state variables of interest into the decisions they ultimately make. The software tool also has the potential to be used in the classroom in the teaching of conditional probabilities, or for health professionals to communicate with one another and their clients about how to reason correctly, albeit ambiguously, from ambiguous empirical evidence. The examples in the paper show that a simple graphical technique can be used to make analysing probabilistic information under ambiguity considerably easier, increasing speed, accuracy, and opening up access to correct statistical reasoning to probabilistically unsophisticated people, even when subject to bounded rationality. A pen and paper version of the software tool is also presented meaning the basic concepts used don't require computer literacy either. What this shows in the end is that while the human brain may have its limitations, its processing power and evolutionary biases (the bounds on human rationality) can be affected through technological innovations which embody our understanding of the way the brain works, and has the potential to be welfare improving through allowing people to make more informed and accurate decisions.

References

- Ballion, A., B. Davies, and P. Wakker. (2010). "Relative Concave Utility for Risk and Ambiguity." Unpublished Working Paper, Econometric Institute, Erasmus University, Rotterdam, the Netherlands.
- Barbey, A. and S. Sloman. (2007). "Base-rate Respect: From Ecological Rationality to Dual Processes." *Behavioral and Brain Sciences*. 30:241-297.
- Bauer, M. (1972). "Bias in Estimates of Conditional Probabilities and Betting Behavior as a Function of Relative Frequency and Validity of Cues in a Cue-Probability Learning Task." *Acta Psychologica*. 36:337-347.
- Birnbaum, M., C. Anderson, and L. Hynan. (1990). "Theories of Bias in Probability Judgement." In *Cognitive Biases (Advances in Psychology)* edited by J. Caverni, J. Fabre, and M. Gonzalez, Chapter 26, pp.477-498. Amsterdam: North Holland.
- Brase, G. (2008). "Frequency Interpretation Of Ambiguous Statistical Information Facilitates Bayesian Reasoning." *Psychometric Bulletin & Review*. 15(2):284-289.
- Burkell, (2004). "What are the Chances? Evaluating Risk and Benefit Information in Consumer Health Materials." *Journal of Medical Library Association*. 92(2):200-208.
- Chapman, G. and J. Liu. (2009). "Numeracy, Frequency, and Bayesian Reasoning." *Judgement and Decision Making*. 4(1):34-40.
- Chen, E. and M. Craske. (1998). "Risk Perceptions and Interpretations of Ambiguity Related to Anxiety During a Stressful Event." *Cognitive Therapy and Research*. 22(2):137-148.
- Cole, W. and J. Davidson. (1989) "New Tools for Analysis and Representation: Graphic Representation Can Lead to Fast and Accurate Bayesian Reasoning." In *Proceedings - Annual Symposium on Computer*

- Applications in Medical Care* edited by L. Kingsland, pp.227-231. Washington:IEEE Computer Society Press.
- Coll, R., J. Coll, and G. Thakur. (1994). "Graphs and Tables: A Four-Factor Experiment." *Communications of the ACM*. 37(4):77-86.
- Dougherty, M. and A. Sprenger. (2006). "The Influence of Improper Sets of Information on Judgement: How Irrelevant Information Can Bias Judged Probability." *Journal of Experimental Psychology: General*. 135(2):262-281.
- Fairman, K. (2006). "Peeking Inside the Statistical Black Box: How to Analyze Quantitative Information and Get it Right the First Time." *Journal of Managed Care Pharmacy*. 13(1):70-74.
- Gigerenzer, G. (2002). *Calculated Risks: How to Know When Numbers Deceive You*. New York: Simon & Schuster.
- Gigerenzer, G. and U. Hoffrage. (1998). "Using Natural Frequencies to Improve Diagnostic Inferences." *Academic Medicine*. 73:538-540.
- Gilboa, I. (2009). *Theory of Decision Under Uncertainty*. Cambridge: Cambridge University Press.
- Kahneman, D., P. Slovic, and A. Tversky, eds. (1982). *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Klibanoff, P., M. Marinacci, and S. Mukerji. (2005). "A Smooth Model of Decision Making Under Ambiguity." *Econometrica*. 73(6):1849-1892.
- Kurzenhauser, S. and R. Hertwig. (2006). "How to Foster Citizens' Statistical Reasoning: Implications for Genetic Counselling." *Community Genetics*. 9:197-203.

- Lad, F. (1996). *Operational Subjective Statistical Methods: A Mathematical, Philosophical, and Historical Introduction*. New York:Wiley-Interscience
- Lewis, C. and G. Keren. (1999). "On the Difficulties Underlying Bayesian Reasoning: A Comment on Gigerenzer and Hoffrage." *Psychological Review*. 106(2):411-416.
- Machina, M. (1987). "Choice Under Uncertainty: Problems Solved and Unsolved." *Journal of Economic Perspectives*. 1(1):121-154.
- Machina, M. (2005). "Choice Under Uncertainty." In *Encyclopaedia of Cognitive Science* edited by L. Nadel, pp.505-514. London: Macmillan.
- Mosleh, A. and V. Bier. (1996). "Uncertainty About Probability: A Reconciliation with the Subjectivist Viewpoint." *IEEE Transactions on Systems, Man, and Cybernetics — Part A: Systems and Humans*. 26(3):303-310.
- Mukerji, S. (2000). "A Survey of Some Applications of the Idea of Ambiguity Aversion in Economics." *International Journal of Approximate Reasoning*. 24:221-234.
- Natter, H. and D. Berry. (2005). "Effects of Information Processing on the Understanding of Risk Information." *Applied Cognitive Psychology*. 19:123-135.
- Nau, R. (2007). "Extensions of the Subjective Expected Utility Model." In *Advances in Decision Analysis: From Foundations to Applications* edited by W. Edwards, R. Miles, and D. von Winterfeldt, Chapter 14, pp.253-278. Cambridge: Cambridge University Press.
- Reyna, V. and C. Brainerd (2008). "Numeracy, Ratio Bias, and Denominator Neglect in Judgements of Risk and Probability." *Learning and Individual Differences*. 18(1):89-107.

- Sedlmeier, P. (2002). "Improving Statistical Reasoning by Using the Right Representational Format." In Proceedings of the Sixth International Conference on Teaching Statistics (ICOTS6) on CD-ROM [ISBN: 0855907827], The Hague: International Statistical Institute.
- Simon, H. (1957). *Models of Man: Social and Rational: Mathematical Essays on Rational Human Behaviour in a Social Setting*. New York: John Wiley & Sons.
- Speier, C. (2006). "The Influence of Information Presentation Formats on Complex Task Decision-Making Performance." *International Journal of Human-Computer Studies*. 64(11):1115–1131.
- Starmer, C. (2000). "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice Under Risk." *Journal of Economic Literature*. 38(2):332-382.
- Stanley, S. (1986). *Air Disasters*. London: Ian Allan.
- Suppes, P. (1957). *Introduction to Logic*. Princeton: D. Van Nostrand Co.
- Thompson, W., E. Schumann. (1987). "Interpretation of Statistical Evidence in Criminal Trials — the Prosecutors Fallacy and the Defense Attorneys Fallacy." *Law and Human Behavior*. 11:167–187.
- Wakker, P. (2004). "On the Composition of Risk Preference and Belief." *Psychological Review*. 111(1):236-241.
- Watkins, S. (2000). "Conviction by Mathematical Error? Doctors and Lawyers Should Get Probability Theory Right." *British Medical Journal*. 320:2-3.