[JEE: Economic Instruction or Content]

UNDERSTANDING BAYES' RULE: INSIGHTS FROM PSYCHOLOGY

August 31, 2000

Abstract

Drawing on insights from psychology, we propose a way to ease the pain of understanding, and teaching, Bayes' Rule.

Introduction

"One of the most famous and useful formulas in probability theory" (Aliprantis and Chakrabarti 2000, p. 91), Bayes' Rule is a difficult concept both to understand and to teach. The apparent lack of intuitive insight into Bayes' Rule has been illustrated in a classic study by Eddy (1982). He provided physicians with various pieces of probabilistic information – the prevalence of breast cancer in a given age group ("base rate"), the sensitivity (the "hit rate" of the test) and specificity (the "false alarm rate" of the test) of mammography tests – and then asked them what the probability was that a randomly drawn woman from the given age group actually would have breast cancer if her mammogram was positive. Inserting Eddy's numbers into Bayes' Rule produces a value of .075. Yet, most of the physicians in Eddy's study estimated the probability to be between .7 and .8. Eddy concluded that the physicians had confused the sensitivity of the test with the posterior probability of the woman having breast cancer. More generally, he argued that "physicians do not manage uncertainty very well, that many physicians make major errors in probabilistic reasoning, and that these errors threaten the quality of medical care." (1982, p. 249)

A similar lack of intuitive insight into Bayes' Rule has been documented in other contexts such as AIDS counselling (i.e., HIV tests, see Holt and Anderson, 1996; Gigerenzer, Hoffrage, and Ebert, 1998) or legal decision making (e.g., Lindsey, Hertwig, and Gigerenzer 1999). People's difficulties in understanding Bayes' Rule have also increasingly been acknowledged in numerous economic contexts. In fact, the very problem has become a staple both in theorizing and experimental tests of economic models (e.g., Camerer 1987; Binmore 1992; Gardner 1995; Holt and Anderson 1996a). Acknowledging people's difficulties, economists have also attempted to find ways to ease the pain of understanding and teaching Bayes' Rule (e.g., Salop 1987 and Holt and Anderson 1996).

In this paper, we present yet another way to facilitate the understanding and teaching of Bayes' Rule. This way is informed by recent results from psychology that demonstrate the importance of how statistical information is represented.

Bayes' Rule and two examples of its application

Bayes' Rule instructs us that in evaluating whether a hypothesis (H) is true relative to its complement (¬H) one ought to incorporate both initial beliefs ("priors") and sample information ("data", or D, from here on) to get updated beliefs ("posteriors"). Formally,

$$\frac{p(H|D)}{p(\neg H|D)} = \frac{p(D|H)\, p(H)}{p(D|\neg H)\, p(\neg H)} \qquad (1)$$

where p(H) and p(¬H) denote the priors, p(H|D) and p(¬H|D) denote the posteriors, and p(D|H)/p(D|¬H) denotes the so-called likelihood ratio. The likelihood ratio is the quotient of what is called the <u>hit rate</u> in the numerator and the <u>false alarm rate</u> in the denominator.[1] Note that

[1] D is sometimes called the individuating, or diagnostic information; the priors are sometimes called the base rates. Unfortunately, although these terms often are used synonymously, the individuating information may not be diagnostic, and priors may not reflect

(1) is the ratio of two other frequently used versions of Bayes' Rule:

(2) $$p(H|D) = \frac{p(D|H)\ p(H)}{p(D)} = \frac{p(D|H)\ p(H)}{p(D|H)\ p(H) + p(D|\neg H)\ p(\neg H)}$$

where (2) is derived from $p(D \cap H) = p(D)\ p(H|D)$ and $p(D \cap H) = p(H)\ p(D|H)$ and involves the conditional probability $p(H|D)$ that one would attach to the event H (or, D) if one knew that the event D had already occurred (Binmore 1992, p. 71). Likewise, $p(D|H)$. Likewise, the derivation of $p(\neg H|D)$.

In what follows, we give two examples of Bayesian calculations. These examples have been proposed to illustrate the importance of Bayes' Rule. However, casual empiricism based on our own teaching experiences suggests that these examples are not easy to understand for the untutored mind. The reason, as we shall show presently, is that the examples are framed in probabilities - a format that is not easily accessible to the human mind.

Example 1. Holt and Anderson (1996, pp. 179/80) propose to motivate the importance of Bayes' Rule by way of the (true) story of a man who was told, following a first-stage test, that he had the virus that caused AIDS, and who committed suicide before follow-up examinations. Pointing out that "the low incidence of the virus in the male population (about 1 in 250 at that time)" and the relatively high false alarm rate of 4 percent combine to a counterintuitive posterior probability of having the virus of about percent, Holt and Anderson (1996) relegate the computation of that number by way of Bayes' Rule to a footnote.

So what is ? Let us summarize the information as given and then apply Bayes' Rule. The prior p(H) is 0.4% ("about one in 250 at that time"), the hit rate p(D/H) is perfect (100%), and the false alarm rate $p(D/\neg H)$ amounts to 4%. If we insert the above values into Equation 2, the probability of an HIV infection given a positive HIV test is

(2) $$p(H|D) = \frac{100\% \times 0.4\%}{100\% \times 0.4\% + 4\% \times 99.6\%} \approx 9\%$$

The posterior probability is thus less than10 %. You knew that, didn't you?

Example 2. Gardner (1995) illustrates how a law firm that just hired a new lawyer might go about assessing her potential. The firm knows from experience that two kinds of lawyers survive its screening process: "Stars" and "ordinary" ones. Star lawyers win 90% of their cases, ordinary ones win 50% of their cases. In addition, the firm knows from experience that only 10% of its newly recruited lawyers turn out to be stars. Preferably that's the kind of lawyer that the firm wants to give a long-term contract to. So, how does the law firm figure out whether it wants to keep a lawyer? Or, in other words, what is the posterior probability of a lawyer being a "star" if she has won her first trial?

---

base rates. We won't address this issue here but see Koehler (1996) for a good discussion.

Once again, let us insert the numerical information -- the prior p(Star) is 10%, the hit rate p(win/Star) is 90% (good but imperfect), and the false alarm rate p(win/ordinary) is 50% -- into Equation 2:

$$p(H|D) = \frac{90\% \times 10\%}{90\% \times 10\% + 50\% \times 90\%} \approx 16.7\ \%$$

The posterior probability that the new lawyer is a star given that she has won her first trial is -- while higher than before -- still relatively low. You knew that too, didn't you?

Present company excepted, it turns out that both undergraduates sitting through tests in psychological laboratories (for a review of these studies see Koehler, 1996), economic laboratories (e.g., Grether 1980, 1992) and experts in medicine (Eddy 1982; Gigerenzer and Hoffrage 1995; Gigerenzer, Hoffrage, and Ebert 1998) and law (Lindsey, Hertwig, and Gigerenzer 1999) have difficulties with Bayesian inference tasks. We are confident that many who tried to teach the Gardner example of the track record principle (a nice example, to be sure) have had similar experiences.

Why is Bayes' Rule so difficult to understand and teach?

Psychologists have given various answers to why Bayes' Rule is so difficult to understand. We consider the two most important ones. The first explanation was proposed in the early 1970's when people became increasingly interested in how people reason about unknown or uncertain aspects of real-world environments. The research program that spurred this interest is the heuristics-and-biases program initiated by Amos Tversky and Daniel Kahneman. This program's strategy has been to measure human decision making against various normative standards taken from probability theory, statistics, and logic. Based on this strategy two major results about people's reasoning under uncertainty emerged: a collection of violations of the normative standards (that in analogy to perceptual illusions are often called "cognitive illusions" or "biases"), and explanations of these illusions in terms of a small number of cognitive heuristics. According to Tversky and Kahneman (1974), people rely on a limited number of heuristics – most prominently representativeness, availability, and anchoring-and-adjustment – that often yield reasonable judgements but sometimes lead to severe and systematic biases.

Concerning Bayesian reasoning, Kahneman and Tversky (1973) proposed that people tend to ignore base rates because they apply the representativeness heuristic. The application of this heuristic asserts that people judge the probability of a sample by assessing "the degree of correspondence [or similarity] between a sample and a population" (Tversky and Kahneman 1983, p. 295). This heuristic can lead to errors because similarity judgements are not always affected by factors that should affect judgements of probability, such as a base rates. From a flurry of laboratory studies that ensued, various researchers concluded that "many (possibly most) subjects generally ignore base rates completely" (Pollard and Evans 1983, p. 124) and that "it appears that people lack the correct programs for many important judgmental tasks" (Slovic, Fischhoff, and Lichtenstein, 1976, p. 174). If these conclusions were indeed correct, then there would not be much hope for improving our understanding of Bayes' Rule. The problem then would not be so much in training, but in our minds, which would lack the correct algorithms and therefore use fallible heuristics. More recently, these conclusions have been challenged on various grounds. In a review of Bayesian reasoning studies, Koehler (1996), for instance argued that "these characteristics of the base rate literature ... are dreadfully misleading." (p. 3), and that "a thorough examination of the base rate literature ... does not support the conventional wisdom that people routinely ignore base rates" (p. 1). He also specified the conditions under which base rates typically are ignored or neglected and the conditions under which they were factored in correctly. His findings are closely linked to the second major reason that psychologists have given to explain why Bayes' Rule is so difficult to understand and teach.

According to this explanation, human cognitive algorithms are not adapted to the format of statistical information that is typically presented in psychological studies. In almost all of the laboratory studies in the heuristics-and-biases tradition, information has been represented in terms of probabilities or percentages – representations of uncertainty that were devised only a few hundred years ago (Gigerenzer et al., 1989). According to psychologists such as Gigerenzer and Hoffrage (1995) and Cosmides and Tooby (1996), people's cognitive algorithms have not caught up (yet). Rather, people remain "intuitive statisticians" that accumulate "natural frequencies" of events, characteristics, etc. in the process of "natural sampling" (rather than systematic sampling).[2] Note that natural frequencies – in contrast to probabilities and percentages that normalize them out – carry information about base rates. If information is presented in normalized values, one has to multiply these by the base rates in order to bring the base rates "back in." Natural frequencies need not be multiplied in this way and they thus make Bayesian computations simpler.

A Bayesian algorithm for computing the posterior probability p(H|D) based on natural frequencies requires solving the following equation:

$$p(H|D) = \frac{d \,\&\, h}{d \,\&\, h + d \,\&\, \neg h}, \tag{3}$$

where d&h (data and hypothesis) is the number of naturally sampled cases with symptom and disease (e.g., positive HIV test and HIV infection), and d&¬h is the number of cases having the symptom but lacking the disease (e.g., positive HIV text and no HIV infection). Note that

---

[2] There are two ways to arrive at frequencies that are not natural frequencies. The first is through systematic sampling, in which the base rates (e.g., 1,000 women with and 1,000 women without breast cancer are tested) are fixed before any observations are made. Thus, these frequencies do not contain information about the base rates of women with and without cancer. A second way to arrive at frequencies that are not natural frequencies is by normalizing natural frequencies with respect to the base rates, that is, by setting the base rates to the same value, such as 1,000. Normalized natural frequencies, like absolute frequencies obtained through systematic sampling, thus have the base-rate information filtered out of them.

Equations (2) and (3) are mathematically equivalent formulations of Bayes' rule. Yet, Equation (3) is computationally simpler, that is, it requires fewer operations (multiplication, addition, or division) to be performed, and the operations can be performed on natural numbers (absolute frequencies) rather than fractions (such as percentages).

While the natural sampling argument stretches credulity for lay people who have to assess the outcome of a first-stage AIDS test, natural sampling may well be a strategy that doctors apply in evaluating uncertain evidence. It is also a plausible procedure that most of us apply in the majority of situations we face daily such as trying to decide whether one of our students' or associates' dismal performance was more than just an aberration. Contrary to the widely accepted statement of the heuristics-and-biases-program that "[i]n his evaluation of evidence, man ... is no Bayesian at all" (Kahneman and Tversky 1972, p. 45), the evidence suggests that "natural sampling" allows us to be good Bayesians if we translate the data in the appropriate format/representation.

Do Frequency Formats Indeed Improve Bayesian Reasoning? Whatever one thinks of the argument that our cognitive algorithms are the baggage of our evolutionary past, psychologists have convincingly demonstrated that translating probabilities (as they are used in the traditional applications of Bayes' Rule) into natural frequencies strongly improves people's insight into Bayesian inference problems (see Koehler 1996 for a comprehensive review). For example, Gigerenzer and Hoffrage (1995) found that with frequency representations, subjects arrived at the numerically exact estimate using a Bayesian algorithm (including pictorial equivalents and shortcuts) in about 50% of the cases. Similarly, Gigerenzer and Hoffrage (1995) found that a frequency format increased the proportion of Bayesian responses in various medical diagnostic tasks from 10% (probability format) to 46%, and Lindsey, Hertwig and Gigerenzer (2000) demonstrated that in legal contexts involving DNA evidence about 40 to 50 percent of a law student sample and 70 to 75 percent of a jurist sample spontaneously derived the correct Bayesian answers using natural frequencies.

In addition, Sedlmeier and Gigerenzer (1999) report several studies that compare learning success for different treatments, including one condition in which the Bayes' Rule in terms of Equation 1 was taught and another condition in which the ability to translate the problem into a frequency representation (rather than a specific rule) was taught. They find that the immediate generalization effect for the representation training was about twice as high as that for rule training. Maybe more importantly, this effect remained stable over follow-up tests (one week, five weeks) whereas performance in the rule-training group showed the typical forgetting curve. Last but not least, Cosmides and Tooby (1996), pushed performance on a Bayesian problem involving medical diagnosis from 36% to 64% (by changing the format of the final question from "What is *the chance* that *a person* who tests positive for the disease will actually have it?" to "*How many people* who test positive for the disease will actually have it?") to 76% (by presenting other elements of the problem description as frequencies) to 92% (by asking subjects to construct a visual representation of the relevant frequencies).

To summarize, psychologists have proposed two major explanations for why people seem to have little insight into Bayesian reasoning. The first is that people lack the cognitive algorithms for dealing with uncertain information, and thus have to rely on heuristics that typically lead to systematic errors. The second explanation suggests that it is the typical representation of uncertainty as probabilities and percentages that causes both lay people and experts to fail at Bayesian inference tasks. The available evidence convincingly demonstrates that when information is presented in terms of frequencies rather than probabilities or percentages, the Bayesian reasoning of both lay people and experts improves significantly. This result has immediate implications for the question of how Bayes' Rule can be taught.

Reframing our running examples in frequencies

As we have seen earlier, the two running examples that we introduced at the beginning, are not easy to compute if one employs probabilities. Following the insights from psychology laid out above, let's see how frequencies help our understanding of the structure of the problem.

Example 1 can be easily translated into natural frequencies. As a matter of fact, Holt and Anderson (1996) did much of the work when they suggested that to explain this point [the low likelihood of having contracted AIDS, not withstanding the first-stage test result] in class, it can be useful to begin with a hypothetical representative group of 1,000 people and ask how likely is it that a person with a positive test actually carries the virus, given an infection rate of one in 250 for the relevant population. On average, only four out of the 1,000 actually have the disease, and the test locates all four of these true positives. However, among the 996 who do not have the disease, the test will falsely identify 4 percent as having it, which is about 40 men (.04 x 996 = 39.84). Hence the test identifies 44 of the 1,000 men as carriers of the virus, four correctly and 40 incorrectly, which means that a positive first-stage actually produces a less-than-10 percent chance of a true positive.
Note that the last sentence is a verbal description of equation (3). Plugging the numbers into this equation yields 4/ (4 + ~40) = 9. Thus, the first-stage indeed produces a less-than-10 percent chance of a true positive. The logic of the Bayesian computation becomes even clearer when information gets translated into a pictorial representation. Figure 1 represents the natural frequencies in terms of a tree. Note that a person does not need to keep track of the whole tree but only of the two pieces of information contained in the bold circles -- these are the hits and false alarms:

[Insert Figure 1 about here]

Example 2, likewise, can also be easily translated into natural frequencies. Recall that we assumed that the firm knows from historical experience that 10 out of 100 lawyers who apply for a job are star lawyers. The firm also knows that 9 out of 10 star lawyers will win their first case while only 45 out of every 90 ordinary lawyers will win their first case. In a new representative sample of lawyers who won the first case, how many of these lawyers can the firm expect to be star lawyers? Plugging the numbers in (2) yields = 9/ (9 + 45) ≈ 16.6. Note that, again, = p(H/D) = f /(f + g) = hits/(hits + false alarms).

The solution to the problem, again, becomes clearer when the frequency interpretation gets translated into a pictorial equivalent.

[Insert Figure 2 about here]

We can see that the likelihood that the new lawyer is a star has increased dramatically after the probationary period. What's reflected in this increase from the prior probability of 10% to the posterior probability of 16.7% is that the binomial distributions over the number of wins evolve differently, with that for ordinarys remaining symmetric and centered on 5 and that for the star becoming skewed and centered roughly on 9 (see Gardner 1995, p. 264).

Conclusion

Understanding and teaching Bayes' Rule is widely acknowledged as a dreaded task by both students and teachers. This is deplorable because understanding the powerful logic of how data can be used to update probabilities is an important skill of the modern citizen. While people such as Kahneman & Tversky have suggested that there may be little hope for people to ever become capable Bayesians, recent results from psychology suggest that this may have been a premature verdict. Rather, it now seems fair to argue that people's cognitive algorithms are not suited for some information formats. If one switches to, or teaches subjects to translate a Bayesian problem into, other information formats – namely, frequency formats - then subjects are Bayesians after all and quite capable to solve the kind of problems that motivated this note. Teaching subjects how to translate a Bayesian problem into a representation that is more suited to her or his cognitive algorithms, has the added bonus that it leads to a skill that is not easily forgotten.

Our proposal to teach subjects representations instead of recipes is not necessarily meant to be a substitute for other approaches. Rather, we see it as a complementary, introductory step to a deeper understanding of Bayes' Rule.

REFERENCES

Aliprantis, C. & S. K. Chakrabarti (2000), *Games and Decision Making*. Oxford University Press.

Binmore, K. (1992). *Fun and Games. A Text on Game Theory*. Lexington, MA: Heath.

Camerer, C.F. (1987). Do Biases in Probability Judgement Matter inMarkets? Experimental Evidence. *American Economic Review* 77 (5), 981-997.

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

Eddy, D.M. (1982). Probabilistic Reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds), *Judgement under uncertainty: Heuristics and biases* (pp. 249 - 267). Cambridge: Cambridge University Press.

Gardner, R. (1995). *Games for Business and Economics.* New York: John Wiley & Sons.

Gigerenzer, G. & U. Hoffrage 1995. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review* 102 (?), 684 - 704.

Gigerenzer, G., Hoffrage, U. & A. Ebert (1998). AIDS Counselling for Low-risk clients. *AIDS Care* 10 (?), 197 - 211.

Gigerenzer, G., Hertwig, R., Hoffrage, U., & P. Sedlmeier (*forthcoming*). Cognitive Illusions Reconsidered. In ...

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., & Krüger, L. (1989). The empire of chance: How probability changed science and everyday life. Cambridge: Cambridge University Press.

Grether, D. M. (1980). Bayes rule as a description model: The representativeness heuristic. *Quarterly Journal of Economics*, 95, 537-557.

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

Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, 73, 538-540.

Holt, C.A. & L. Anderson 1996. Understanding Bayes' Rule. *Journal of Economic Perspectives* 10 (2), 179 - 183.

Holt, C.A. & L. Anderson 1996a. Information Cascades. *Journal of Economic Perspectives* 10 (4), 187- 193.

Kahneman, D. & A. Tversky (1972). Subjective probability: a judgement of representativeness. *Cognitive Psychology* 3(?), 430-454.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.

Koehler, J.J. (1996). The base rate fallacy considered: Descriptive, normative, and methodological challenges. *The Behavioral and Brain Sciences* 19, 1 - 53

Lindsey, S., R. Hertwig, & G. Gigerenzer (2000). Communicating statistical evidence. Manuscript submitted for publication.

Pollard, P. & J. St. Evans (1983). The role of 'representativeness' in statistical inference: A critical appraisal. In: *Thinking And Reasoning: Psychological Approaches*, ed. J. St. B. T. Evans. Routledge/Kegan Paul.

Salop, S. 1987. Evaluating Uncertain Evidence With Sir Thomas Bayes: A Note for Teachers. *Journal of Economic Perspectives* 1 (1), 155 - 160.

Sedlmeier, P. & G. Gigerenzer 1999. Teaching Bayesian reasoning in less than two hours. Manuscript submitted for publication.

Slovic, P. Fischhoff, B. & Lichtenstein, S. (1976), Cognitive processes and societal risk taking. In J.S. Carril & J.W. Payne (eds), *Cognition and Social Behavior*. Erlbaum.

Tversky, A. & Kahneman, D. & (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124 - 1131.

Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probabilistic judgement. *Psychological Review*, 90, 293-315.