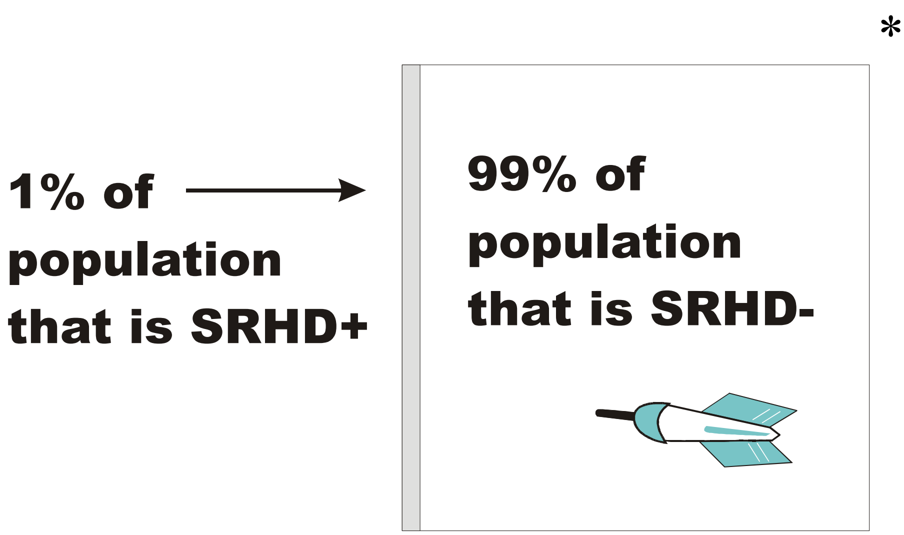
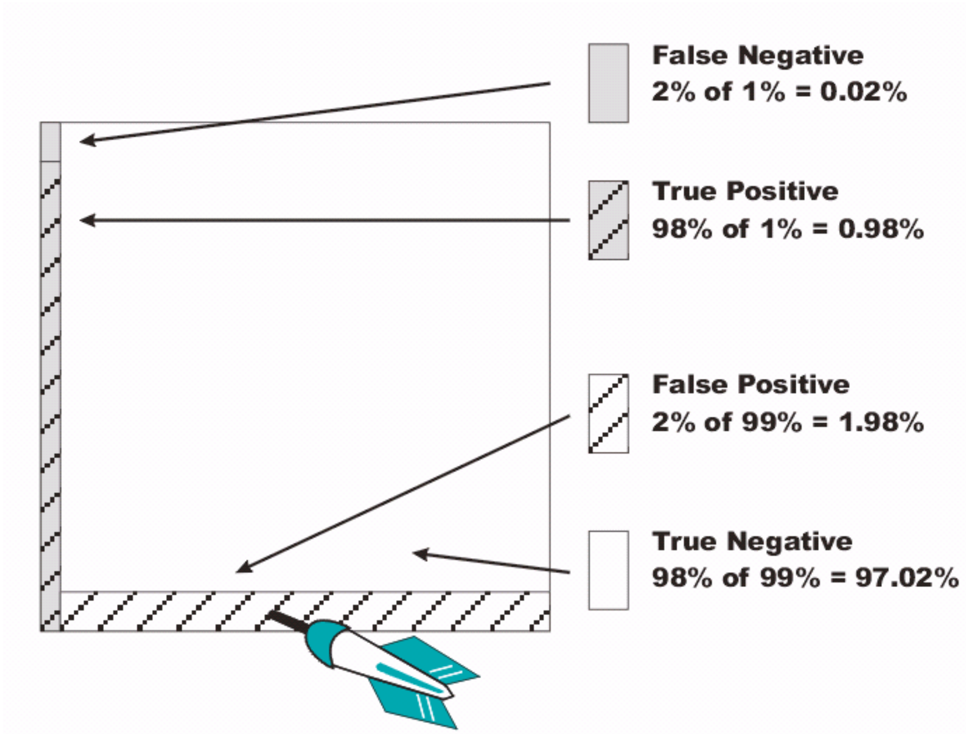
Bayesian Thinking

Bayesian Thinking (based on Bayes Theorem) can provide unintuitive insights in a wide variety of situations. As a medical example, consider some rare horrible disease (SRHD), which strikes one person in a hundred independent of ethnicity, geographical area, or socioeconomic status. Suppose that there is a 98% effective diagnostic test for SRHD. That is, 98% of people who are infected display a positive test result, while 98% of those not infected display a negative test result. Now imagine someone, who in routine screening, has just tested positive for SRHD. What is the chance they are ***actually*** infected?

The gut response is a 98% chance of infection. But Bayesian Analysis says it's only one out of three! Here's why. Imagine throwing a dart at the entire population in question. Then only a 1% sliver (shaded area) represents those infected (SRHD+).



Now consider the testing. A positive test (hatched area) can occur in two ways: true positive (98% of 1%) or false positive (2% of 99%). So the real question is "what is the chance of hitting a shaded area given that you know you have hit a hatched area." Clearly about one third.



This is a strong argument against universal HIV testing, as the majority of positive tests would be false. And there is an exact analogy with homeland security screening. Suppose neuro-scientists develop a brain wave detector that has a 99% chance of detecting someone who posed a serious terrorist threat. Let's assume there are currently 3,000 such people in the country. Then if you tested people at random, the chance that someone who triggered the device was a true terrorist would be only about one in a thousand.

Two more classic examples of Baysian Analysis come from WWII. The first involves the statistician Abraham Wald, who was asked to determine where to add a limited amount of armor plating on bombers to protect them against anti-aircraft fire. When a plane returned from a mission the Air Force would record the location of all bullet holes, so there was plenty of data. Furthermore, there were clearly more hits in some parts of the planes than others, so the answer was obvious: put armor plating only where there was a high density of bullet holes. Obvious, but wrong. Wald realized that he was only looking at data from planes that survived their mission, implying that these bullet holes didn't matter. The correct solution was to put amour where there were ***no*** bullet holes.

A second story relates to placing anti-aircraft guns and gunnery crews aboard merchant ships to protect them against enemy bombers. Since both the weapons and crew training were in scarce supply, a handful of vessels were equipped to test the idea for several weeks. At the end of the trial period the gunnery crews were asked how many planes they had shot down, and the answer, unfortunately, was "none." The entire plan was about to be scrapped, when someone thought to ask a different question: "how many of the ships in the test were sunk by bombers?" As it turned out the answer was again "none," which was much lower than would have been expected otherwise. Apparently the enemy pilots did not enjoy flying through tracer fire, and the program was a success.

So why hasn't the concept of Bayesian Analysis been applied more widely? Because concepts are hard to apply until they evolve into tools. Consider the concept of transferring momentum from a heavy block of iron into a thin metal spike, thereby forcing it through two adjacent pieces of wood to hold them together. I myself prefer to use a hammer, a tool from an earlier age based on the laws of physics. Today, we are at the dawn of the Information Age, in which tools based on mathematics are allowing us to apply concepts such as Bayesian Analysis without breaking a sweat. These new information tools are not guaranteed to give us the right answer, but they often provide more valuable, by helping us ask the right question.

* Drawings from Dr. Savage's Decision Making with Insight, - Text and Software, Duxbury Press, Belmont CA 2003. Not to be duplicated without permission.