DOCID: 3927956

The Apparent Paradox of Bayes Factors

BY F. T. LEAHY

Unclassified

There are, let us assume, two hypotheses that have associated with them a priori odds. That is, we state, for example, that one hypothesis is twice as likely to be correct as the other.

The top card of a well-shuffled face-down deck has a priori odds of 8 to 5 of being a small card (i.e., duece through the ninespot) rather than an honor. This type of statement cannot be proved, but most people accept it as self-evident.

Let us perform some kind of experiment which will result in our obtaining a "Bayes Factor"—a process to be described presently. We will then have a posteriori odds which can be multiplied by the a priori odds to obtain the actual odds that apply to the hypotheses.

(*Note:* When computers are being utilized, and frequently otherwise, the *logarithm* of the a priori odds and the *logarithm* of the Bayes Factor, which incidentally is synonymous with a posteriori odds, are added together to obtain the *logarithm* of the actual, or true, odds.)

A Bayes Factor is obtained by determining the ratio of the probabilities of observing various phenomena, or characteristics (often called "events"), under each of the two hypotheses, to be called I and II.

A card from a pinochle deck has a probability of 1/6 of being an ace; a card from an ordinary deck, a probability of 1/13 of being an ace.

The ratio of these probabilities, 13/6, will be used when we actually observe an ace, upon the facing of a concealed card, if we are to obtain a Bayes Factor in favor of the hypothesis that the card came from a pinochle deck (I), rather than the hypothesis that the card came from an ordinary deck (II).

The paradox seems to arise when we ask whether Bayes Factors are useful when the a priori odds are unknown. Some statisticians and philosophers assert that "nothing" can be known about actual odds if nothing is known about the a priori odds. For, they ask, what is the magnitude of two (the assumed a posteriori odds) times an unknown number? (The a priori odds, like the a posteriori, can assume any positive numerical value—or zero—fractional or whole.)

Approved for release by NSA on 11-29-2011, Transparency Case# 63852

"Well, at least in such a case the true odds are twice as big as they were originally," we respond.

"With that statement we do agree, "reply the philosophers, "but you haven't ruled out any specific value for the true odds." And naturally we admit that it is only the true odds that interest us.

The philosophers' argument is demonstrably false, but the fallacy is not readily recognizable. I have posed this problem to many serious thinkers, but I have seldom, if ever, received a satisfactory explanation.

In cryptanalysis, we frequently perform a million or more consecutive experiments, with a Bayes Factor computed for each experiment. For example, for a particular experiment, the two hypotheses might be: (1) This specific message was enciphered (on a given device) with an initial window setting of ABCDEF. (II) This message was *not* enciphered with this window setting (and conceivably not even on the device assumed). Hypothesis II is called the null hypothesis, and hence all the characters of the message are assumed to be equiprobable.

When the computer (or analytic device) locates and prints out a setting and the deciphered text of the message, it is (or could be) accompanied by the Bayes Factor (i.e., the a posteriori odds) in favor of Hypothesis I.

Let us assume that they are 10^{12} to 1. But how do we know that there were not a priori odds of 10^{18} to 1 against the choice* of this specific window setting? Or, to consider another situation, let us assume that there is a different window setting with a Bayes Factor of 10^{2} to 1 in its favor. How do we know that the latter setting did not have a priori odds of 10^{11} to 1 over the former setting? For, if this had been true, the latter setting would indeed be ten times as likely to be correct as the former. Yet the computer has stopped and printed out the former setting as the (only) correct one. (I might add that no one has yet disputed the assertion that the text printed out is *right*.)

Some individuals have said that there is no mystery or fallacy to be explained. Some say that we have merely ordered all our answers in accordance with decreasing a posteriori odds, and this is, of course, true.

Others assert that we have *assumed* all settings to have equal a priori probabilities. And finally, some others have pointed out that there is only a finite number of settings that exist, of which one must be the correct one.

*Or even *infinite* odds against; i.e., if this setting were *forbidden*.

In my opinion, none of the explanations can be accepted as throwing much light on the paradox.

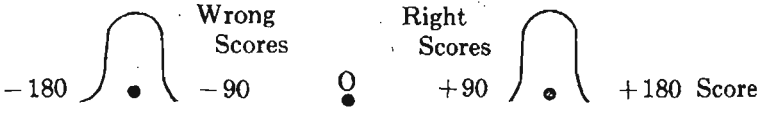
The statement that we have tacitly assumed all settings to have equal a priori probabilities cannot be accepted. We may know that some 50% of the settings have a priori probabilities of zero; i.e., their selection is forbidden because of some weird set of rules imposed upon the cryptographer from above. Or to illustrate this point even more strongly, we can set up a laboratory experiment in which the most eccentric rules imaginable can be created to govern the probabilities involved in the selection of the window settings. For example, every setting might be one-millionth as likely as its right-hand neighbor. Or, there might be an arbitrarily large or small subset of the settings that have been assigned an unconditional a priori probability of zero; that is, under no circumstances whatsoever may a member of this subset be selected.

Nevertheless, none of these suggested stratagems will prevent our computer (or analytic device) from arriving at the correct answer, nor can they even increase its difficulties in doing so.

Hence, it certainly appears that knowledge, however slight, of the a priori odds is unnecessary, nor are "lucky guesses" of importance. For there can exist for the cryptographer no assignment of a priori odds (whether ingenious or otherwise) that can adversely affect the usefulness of our computer program.

The above paragraph is the most important of this entire article since it appears to be contradicted in many of the books on Statistics that condescend to discuss Bayes Factors. For example, the Encyclopedia Brittanica considers the "Achilles' Heel" of Bayes Factors to be their dependence upon a priori probabilities. Some state that if the a priori probabilities are not known, Bayes' Theorem cannot be employed.

Now I'd like to offer my explanation as to why the a priori odds do *not* make any difference in most cryptanalytical applications of Bayes Factors. When planning experiments (such as one whose purpose is to find a correct window setting), we are able first to compute the mean and variance of the scores (i.e., the log Bayes Factors) of both the right and wrong answers. The distribution curves of the scores are approzimately normal in practice, but it is not necessary to insist that they have to be.



A window setting that *in fact* was unselected, whether its a priori probability of being chosen actually was .0000 or .9999, will, during the course of our experiment, yield a score (in this illustration) of somewhere between −180 and −90.

No scores will ordinarily lie between −90 and +90; that is, we "expect" no answers in this region. (The mathematical expectation might be "one-hundredth of an answer.")

The right answer that perforce *was* selected might have been associated with any non-zero a priori probability, no matter how large or small. But it will score between +90 and +180. The computer, of course, had previously been instructed to stop and decipher *any* message that scored over +90. Thus, only right answers will ever actually be printed. And the a priori odds have been completely ignored!

Final comments: It is important to point out that in any experiments where the domain of wrong answers overlaps the domain of right answers, the a posteriori odds could begin to assume importance. I am not overlooking the fact that two normal distributions always have *some* overlap, but in certain regions this can be insignificantly small. To assume that there exists an empty region between two such distributions, of course, raises no practical difficulties.