Statistical Science
1987, Vol. 2, No. 3, 211–212

In This Issue

The following type of example is often presented in introductory probability and statistics courses to help sharpen students' intuition about the importance of background rates in calculating probabilities: Suppose that you are walking down the street and notice that the Department of Public Health is giving a free medical test for a certain rare disease. The test is 90% reliable in the following sense: If a person has the disease, there is a probability of 0.9 that the test will give a positive response (the "sensitivity" of the test); and if a person does not have the disease, there is a probability of 0.9 that the test will give a negative response (the "specificity" of the test). Data indicate that your chances of having the disease are only 1 in 5000. However, because the test costs you nothing (you have already paid for it with your taxes), and it is fast and harmless, you decide to stop and take the test. A few days later you learn that you had a positive response to the test. What is now the probability that you have the disease?

Many beginning students feel that this probability should be about 0.9, but that feeling mistakenly ignores the small prior probability of 0.0002 that you had the disease. The correct posterior probability is found by Bayes theorem to be 0.0018. Your probability of having the disease is now 9 times as large as it was before you took the test, but it is still extremely small. The intuitive explanation is that because the test has a 10% rate of producing false positives, there will be about 500 positive responses among a group of 5000 persons, but on the average only one person in the group will have the disease.

It is this large number of false positives that has led various interested parties to question the effectiveness of large-scale medical screening tests for populations in which the prevalence of the disease is low, and which is the subject of the opening article by Joseph L. Gastwirth in this issue. He considers problems in which the prevalence, as well as the sensitivity and the specificity of the test, are unknown, and discusses effective experimental designs for estimating these quantities in order to obtain an estimate of the posterior probability given a positive response that will have small variance. He describes two applications that have been very much in the news in recent years: the screening of general populations for the presence of antibodies to the AIDS virus and the screening of the employees, or potential employees, of an organization with polygraph (or "lie detector") tests.

In his discussion of this article, D. H. Kaye considers the standards that are used for the admissibility of polygraph evidence in court, and the relevance of Gastwirth's work to the legal question of admissibility. John C. Kircher and David C. Raskin point out that the problem of low base rates has been discussed for many years in the psychology literature, and describe the many different contexts in which polygraph tests are used. Janet Wittes emphasizes that the context of a medical screening determines whether the sensitivity or the specificity of the test is more important. Judith D. Goldberg points out that not only prevalence, but also false positive and false negative rates, can vary from group to group. Seymour Geisser sketches a Bayesian predictive approach to the problems addressed by Gastwirth. Finally, Beth C. Gladen comments that in many situations the application of a confirmatory test following a positive response would make variance calculations relatively unimportant.

* * *

In his article, "Uncertainty, policy analysis, and statistics," James S. Hodges states that "No existing school of statistical thinking provides a comprehensive framework for considering the various types of uncertainty and the tradeoffs among them that analysts must make." He describes three major types of uncertainty: (1) structural uncertainty, which is uncertainty about the model that is used; (2) risk, which is uncertainty due to statistical or stochastic variability given the model; and (3) technical uncertainty, which is uncertainty due to data processing and the use of approximations. He argues that the absence of a system that properly accounts for all these types "creates an inherent tendency for analyses to understate uncertainty about predictions . . . which can lead to invisible biases in policy considerations." He believes that the de Finetti approach comes closest to providing such a system, and he tries in this paper to develop further the connection between that approach and real policy applications.

In his comment, David Freedman states that "Good statistical analysis can be done in either the frequentist or the Bayesian framework. However, for either approach to succeed, the analyst has to get the model right, or close enough." Seymour Geisser points out that there is a fundamental principle in the de Finetti approach to statistics "that statisticians (even Bayesian predictivists) often ignore." Peter J. Huber comments on "the problem of the infinite regress, and the question of whether and when to combine different kinds of uncertainty." Joseph B. Kadane stresses three important aspects of de Finetti's approach: the subjectivity of probability, the emphasis on prevision and