Épreuve de section européenne

False positive paradox

When someone takes a medical test, the result is said to be:

- a *true positive* if the test is positive and the person is infected;
- a *true negative* if the test is negative and the person is healthy;
- a *false positive* if the test is positive but the person is healthy;
- a *false negative* if the test is negative but the person is infected.

The false positive paradox is a statistical result where false positive tests are more probable than true positive tests, occurring when the overall population has a low incidence of a condition and the incidence rate is lower than the false positive rate.

The probability of a positive test result is determined not only by the accuracy of the test but by the characteristics of the sampled population. When the incidence (the proportion of those who have a given condition) is lower than the test's false positive rate, even tests that have a very low chance of giving a false positive in an individual case will give more false than true positives overall.

So, in a society with very few infected people – fewer proportionately than the test gives false positives – there will actually be more who test positive for a disease incorrectly and don't have it than those who test positive accurately and do. The paradox has surprised many.

Imagine running an HIV test on a population A, in which 1 person in 10,000 is infected. The test has a false positive rate of 0.0004 and no false negative rate. The expected outcome of a million tests on this population would be:

- 100 people would receive a true positive;
- 400 people would receive a false positive.

Only 100 of the 500 total people with a positive test result are actually infected. So, the probability of being infected after you are told the test is positive is only 20% for a test that otherwise appears to be "over 99.95% accurate".

Adapted from *"False positive paradox"*, Wikipedia

Questions

1. Explain with your own words the difference between a true and false positive test. Which of the two results is considered an error?
2. In the above example, explain the numbers 100 and 400 of people receiving true and false positive tests.
3. Why is the test said to be *"over 99.95%"* accurate?
4. Imagine you run a test having the same characteristics on a population B in which 200 out of 10,000 are infected. What is the expected outcome of a million tests on this population, and what is the probability of being infected after you are told the test is positive?
5. Explain the sentence *"The probability of a positive test result is determined not only by the accuracy of the test but by the characteristics of the sampled population"*.