BACCALAUREAT-DNL Mathématiques/Anglais
Session 2018

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.

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

Tasks:

1. Explain in your own words the difference between a true and false positive test. Which of those two events is considered an error?
2. In the example, explain how to get the numbers 100 and 400 people receiving true and false positive tests. Is this situation a false positive paradox?
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? Assuming your test is positive, determine the probability of being infected.
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.”*

Elements de correction

1. The false positive is an error : the test is wrongly positive because the person is healthy.

2. The incidence rate is 1 out of 10,000 so you get 100 infected people over a million.
   The false positive rate is 0.0004 so you get 400 false positive over a million tests.
   We have more false positive than true positive : it’s a false positive paradox.

3. There is no false negative : the only error could be a false positive with a 0.0004 rate, so less than 0.05 %. It means the accuracy rate of the test is more than 99.95 %.

4. The test has not changed, so you still get 400 false positive.
   The incidence of the disease is 200 out of 10,000 so you get 20,000 infected people.
   If your test is positive, the probability to be infected is 20000/20400 ≈ 0,98.

5. The two previous examples show that the characteristics of the population (few or many people infected) will change the probability to get a positive result. The accuracy of the test only change the number of false positive results.