Sensitivity and Specificity

The words "sensitivity" and "specificity" have their origins in screening tests for diseases. When a single test is performed, the person may in fact have the disease or the person may be disease free. The test result may be positive, indicating the presence of disease, or the test result may be negative, indicating the absence of the disease. The table below displays test results in the columns and true status of the person being tested in the rows.

		Test Result (T)	
		Positive $(+)$	Negative $(-)$
True Status of Nature (S)	Disease $(+)$	a	b
	No Disease $(-)$	c	d

Though these tests are generally quite accurate, they still make errors that we need to account for.

Sensitivity: We define sensitivity as the probability that the test says a person has the disease when in fact they do have the disease. This is $P\left(T^{+} \mid S^{+}\right)=\frac{a}{a+b}$. Sensitivity is a measure of how likely it is for a test to pick up the presence of a disease in a person who has it.

Specificity: We define specificity as the probability that the test says a person does not have the disease when in fact they are disease free. This is $P\left(T^{-} \mid S^{-}\right)=\frac{d}{c+d}$.

Ideally, a test should have high sensitivity and high specificity. Sometimes there are tradeoffs in terms of sensitivity and specificity. For example, we can make a test have very high sensitivity, but this sometimes results in low specificity. Generally we are able to keep both sensitivity and specificity high in screening tests, but we still get false positives and false negatives.

False Positive: A false positive occurs when the test reports a positive result for a person who is disease free. The false positive rate is given by $P\left(S^{-} \mid T^{+}\right)=\frac{c}{a+c}$. Ideally we would like the value of c to be zero; however, this is generally impossible to achieve in a screening test involving a large population.

False Negative: A false negative occurs when the test reports a negative result for a person who actually has the disease. The false negative rate is given by: $P\left(S^{+} \mid T^{-}\right)=\frac{b}{b+d}$.

Which false result is the more serious depends on the situation. But we generally worry more about false positives in screening tests. We don't want to tell someone that they have a serious disease when they do not really have it.

The following example is take from *Statistical Methods for the Social Sciences, 3rd edition* (Alan Agresti and Barbara Finlay, Prentice Hall, page 287). The data are provided for the results from a screening test for HIV that was performed on a group of 100,000 people. Note that the prevalence rate of HIV in this group was very low.

		Test Result		
		Positive $(+)$	Negative $(-)$	
HIV Status	Positive $(+)$	475	25	500
	Negative $(-)$	4975	94525	99500
		5450	94550	100000

Based on the results above, the sensitivity of this test was $\frac{475}{500} = 0.95$ and the specificity was $\frac{94525}{99500} = 0.95$. This test appears to be pretty good. What are the false positive and false negative rates for this test?

False positive rate = $\frac{4975}{5450} = 0.91$ False negative rate = $\frac{25}{94550} = 0.0003$

Even when a test has a high sensitivity and specificity, you are still going to get a high false positive rate if you are screening a large population where the prevalence of the disease is low. False positive rate is not just a function of sensitivity and specificity. It is also a function of the prevalence rate of the disease in the population you are testing. Thus there is danger in indiscriminately applying a screening test to a large population where the prevalence rate of the disease is very low.

In a significance test, sensitivity and specificity relate to correct decisions and false positives and false negatives correspond to Type I and Type II errors. Consider the columns in the table below. A positive test result corresponds to a statistically significant test result and thus leads to rejection of the null hypothesis. A negative test result corresponds to a test result that is not statistically significant; thus we decide not to reject H_o . Now consider the rows in this table. If the true status of nature is positive (that is, if something is going on and what we are testing for is really true), then the alternative hypothesis is true. if what we are testing for is not true (and nothing is going on), then the null hypothesis is correct.

		Test Result (T)	
		Positive $(+)$	Negative $(-)$
True Status of Nature (S)	H_a True $(+)$	a	b (Type II)
	H_0 "True" $(-)$	c (Type I)	d

Note that the upper left entry in the table corresponds to the correct decision to reject the null hypothesis when the alternative is really true. A high value of *a* results in a high probability of detecting that a null hypothesis is false and corresponds to a test that is very sensitive. The lower right entry corresponds to the correct decision not to reject the null hypothesis when it should not be rejected. A high value of *d* results in a high probability of not rejecting a "true" null hypothesis and corresponds to a test with high specificity. The lower left entry corresponds to a test result that causes us to reject the null hypothesis when it is really "true;" this entry reflects Type I error and corresponds to a false positive. The upper right entry corresponds to a test result that causes us not to reject the null hypothesis when it is not true; this entry reflects Type II error and corresponds to a false negative. Note that a small value of *c* is consistent with low probability of Type I error and a low value of *b* is consistent with low probability of Type II error.