TP, TN, FP, FN and Derived Measures for a Test

The following is a C*ontingency Table* of the types of errors and success (or hits) in a test for the presence of some anomaly (a tumor, a pathology, etc,), in general for the output of a binary classifier, a decision process or a diagnostic procedure:

Case	**Classifier/Decision or test outcome**	**Reality (Truth)**
False Positive (**FP**)	Anomaly	Normal
False Negative (**FN**)	Normal	Anomaly
True Positive (**TP**)	Anomaly	Anomaly
True Negative (**TN**)	Normal	Normal
$\text{Sensitivity} = \dfrac{\text{TP}}{\text{TP}+\text{FN}}\times 100\%$		

The test outcome (middle column) can be also stated in general as "positive/negative" (it would be confusing to use "true/false"). The following measures qualify the performance of a test, classifier or decision process (a test in biology, or any science; in manufacturing or defect inspection) and also are used in medicine to assess the accuracy of a diagnostic procedure:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN}+\text{FP}}\times 100\%$$

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP+TN}+\text{FP+FN}}\times 100\%$$

$$\text{Prevalence} = \frac{\text{TP}+\text{FN}}{\text{TP+TN}+\text{FP+FN}}\times 100\%$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}}\times 100\%$$

The ***sensitivity*** (or *true positive rate*) of a test is the probability (*a posteriori*) of its yielding ***true-positive*** (TP) results in patients who actually have the disease. A test with high sensitivity has a low ***false-negative*** (FN) rate. The quantity (1–*Sensitivity*) is known as ***false negative rate.*** In other contexts, more general than binary classification, *sensitivity* is also known as *recall.*

The ***specificity*** (or ***true negative rate***) of a test is the probability (*a posteriori*) of its yielding negative results in who do not have the disease. A test with high specificity has a low ***false-positive*** (FP) rate; it does not give a FP result in many patients who do not have the disease. *FP is known as a Type-I error and FN a Type-II error. The quantity (1-*Specificity*) is known as ***false positive rate.***

The ***accuracy*** is similar to the ***prior probability*** or ***prevalence,*** of the condition prior to the test (all diseased persons divided by all persons). Note that:

$$accuracy = (sensitivity)(prevalence) + (specificity)(1 - prevalence)$$

The ***Precision*** can be seen as a measure of "exactness" or fidelity, whereas ***recall*** (or *sensitivity*, in binary classification) is a measure of completeness.

*Other factors

(**Note**: in order to interpret as probabilities, it is common to normalize to fractions in [0,1], and not to express the measures in percentages (hence, "×100%" does not appear).

There are also, especially in medical diagnosis, the following measures:

The ***Likelihood ratio positive*** = *sensitivity* / (1 − *specificity*)

The ***Likelihood ratio negative*** = (1 − *sensitivity*) / *specificity*

The ***False positive rate*** **(**α**)** = 1 − *specificity* = FP / (FP + TN)

The ***False negative rate*** **(**β**)** = 1 − *sensitivity* = FN / (TP + FN)

The sensitivity is also known as *power* and equals 1 − β.

The ***Positive Predictive Value***, also known as the ***precision rate***, or the ***post-test probability*** of a disease. It is the proportion of patients with positive test results who are correctly diagnosed. It is the most important measure of a diagnostic method as it reflects the probability that a positive test reflects the underlying condition being tested for. Its value does however depend on the prevalence of the disease, which may vary. In terms of the latter definitions, we have:

$$\text{PPV} = \frac{(sensitivity)(prevalence)}{(sensitivity)(prevalence) + (1 - specificity)(1 - prevalence)} = \frac{\text{TP}}{\text{TP+FP}}$$

and, similarly, we have the ***Negative Predictive Value***:

$$\text{NPV} = \frac{\text{TN}}{\text{TN+FN}}$$

***Exercise:** write NPV in terms of sensitivity, prevalence, specificity, etc and/or PPV.

§§