Confusion matrix and evaluation metrics

Justin

# Confusion matrix

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Actuality** | |
|  |  | **P** | **N** |
| **Prediction** | **P** | TP (hit) | FP (false alarm, type I error) |
| **N** | FN (miss, type II error) | TN (correct rejection) |

Total is the sum of TP, FP, FN and FN

Sensitivity, hit rate, recall, TPR (no miss → high recall):

Precision (no false alarm → high precision):

Accuracy:

FPR:

F1 score:

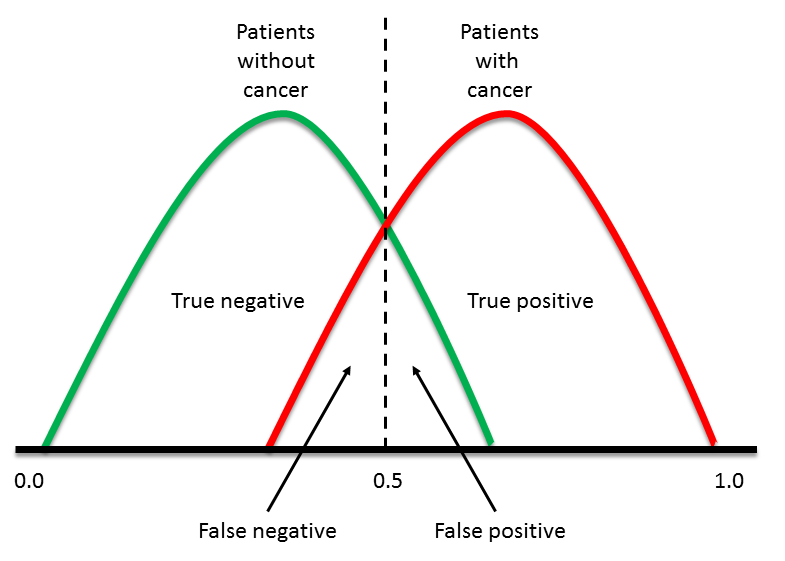


Figure . Just one way to visualize the confusion matrix

# Errors

## Type I error (false positive, false alarm)

Type I error occurs when **rejecting the null hypothesis when it’s true**. The type I error rate or significant level is usually set to 5%, implying that it is acceptable to have a 5% probability of incorrectly rejecting the null hypothesis.

## Type II error (miss)

Type II error occurs when **failing to reject the null hypothesis when it’s false**. The type II error rate is denoted by , and is related to the power of a test (which is ).

# Receiver operating characteristic (ROC, TPR vs FPR)

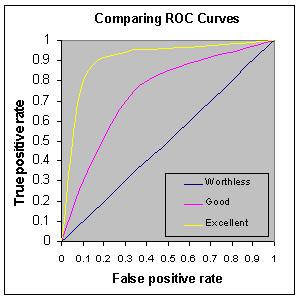


Figure . Note when TPR == FPR, the prediction is worthless!