

ICT for Health Laboratory # 3 ROC - Covid-19 serological tests

Monica Visintin

Politecnico di Torino



November 20th 2023

Table of Contents

Description of laboratory # 3

Data analysis

Sensitivity and specificity

ROC

Threshold setting

Description [1]

Covid-19 can be diagnosed using naso-pharyngeal swabs or using serological tests (blood test). Serological tests, at least at the beginning, were faster, but naso-pharyngeal swabs were more reliable (higher sensitivity and specificity).

In the experiment, naso-pharyngeal and **two serological tests** were used on some subjects; for the serological tests the marker levels (IGg) were recorded, for the swab test only the result (positive/negative) was recorded.

Goal of the lab is to **set the threshold of the marker above which each serological test is considered positive, using ROC curves**, and compare the two serological tests, deciding which should be used.

Table of Contents

Description of laboratory # 3

Data analysis

Sensitivity and specificity

ROC

Threshold setting

Data analysis

The input file is `covid_serological_results.csv` and contains 3 columns:

- ▶ `COVID_swab_res` which is the swab result: 0 means negative, 2 means positive, 1 means uncertain result
- ▶ `IgG_test1_titre` which is the level of Immunoglobulin G (IgG) in test #1; values range from 2.5 to 314
- ▶ `IgG_test2_titre` which is the level of Immunoglobulin G (IgG) in test #2; values range from 0 to 9.71

The dataset includes 879 rows, without missing values. Note that the scale with which IgG is measured is somehow arbitrary and therefore each test will have its own threshold.

We will consider swab test result as ground truth (which might not be correct, since also the swab test has its own sensitivity and specificity), and we will **drop patients with uncertain (i.e. equal to 1) swab tests**.

Read the data [1]

The first lines in the script are the following:

```
1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 import sklearn.cluster as sk
5
6 plt.close('all')
7 xx=pd.read_csv("covid_serological_results.csv")
8 xx=xx[xx.COVID_swab_res!=1]# remove rows with swab test = 1
9 xx.COVID_swab_res[xx.COVID_swab_res==2]=1# replace swab test = 2 with swab test = 1
10 swab=xx.COVID_swab_res.values
11 Test1=xx.IgG_Test1_titre.values
12 Test2=xx.IgG_Test2_titre.values
```

Read the data [2]

Perform the usual data analysis using `xx.describe()`, `pd.plotting.scatter_matrix()`, etc.

Table of Contents

Description of laboratory # 3

Data analysis

Sensitivity and specificity

ROC

Threshold setting

Find sensitivity and specificity [1]

Let us focus on Test2 (**you have to repeat everything for Test1**). We arbitrarily set the threshold to 5, and we find sensitivity and specificity for this threshold.

```
10 x=Test2
11 y=swab
12 x0=x[swab==0] # test results for healthy patients
13 x1=x[swab==1] # test results for ill patients
14 Np=np.sum(swab==1) # number of ill patients
15 Nn=np.sum(swab==0) # number of healthy patients
16 thresh = 5 # example of threshold
17 n1=np.sum(x1>thresh) # number of true positives for the given thresh
18 sens=n1/Np # sensitivity
19 n0=np.sum(x0<thresh) # number of true negatives
20 spec=n0/Nn # specificity
```

Sensitivity is 0.48 (not nice) and specificity is 0.96 (good).

↓
P(Tp|D)

Find sensitivity and specificity [2]

Write a **method (def)** that generates the values of sensitivity and specificity versus the threshold, considering a reasonable interval for the threshold. A good choice of the set of thresholds is made of the sorted values of Test2; otherwise use `np.arange(start, stop, step)`.

Remember to include **threshold zero**.

The method should output thresholds, sensitivity and specificity.

Plot sensitivity and specificity versus the threshold (two curves in the same graph).

Find sensitivity and specificity [3]

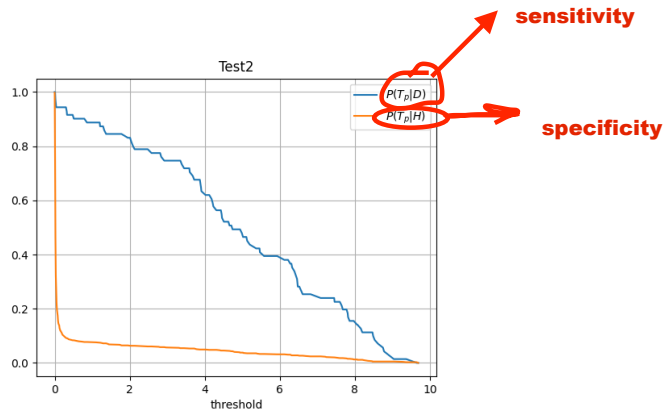


Table of Contents

Description of laboratory # 3

Data analysis

Sensitivity and specificity

ROC

Threshold setting

ROC [1]

ROC (Receiver Operating Curve) is sensitivity $P(T_p|D)$ (y-axis) versus false positive rate $P(T_p|H)$ (x-axis):

$$P(T_p|D) \text{ versus } P(T_p|H)$$

Therefore, to plot ROC, we are not interested in specificity, but in false positive rate, which is 1 minus specificity:

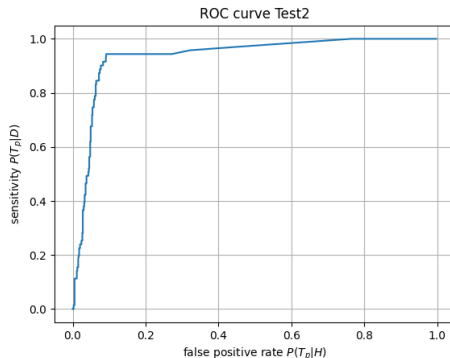
$$P(T_p|H) = 1 - P(T_n|H)$$

21 FA=1-spec

Generate the ROC curve by plotting sensitivity versus false positive rate. Set grid, the x and y labels, etc.

ROC [2]

The ROC curve will be similar to the following:



The ideal working point is for false positive rate equal to 0 and sensitivity equal to 1, which is not reached by this test; however the results are not so bad, after all.

ROC [3]

evaluate the area under the roc curve to understand which test is better

A parameter that is typically used to understand if a test is better than another is the **area under the ROC curve (AUC)**: **the higher the area, the better the test.**

Write a method to find the area under a curve (input vectors x and y). Use the **trapezoidal rule** to evaluate the area.

Check

Use methods available in Scikit Learn to check your results:

```
1 from sklearn import metrics
2 from sklearn.metrics import roc_auc_score, average_precision_score
3 fpr, tpr, thresh=metrics.roc_curve(swab, Test2, pos_label=1)
4 AUC_Test2=roc_auc_score(swab, Test2)
```


Table of Contents

Description of laboratory # 3

Data analysis

Sensitivity and specificity

ROC

Threshold setting

How to set the threshold [1]

There are several ways to set the threshold above which the test is considered positive:

1. The required sensitivity is set, and the associated false positive rate (and therefore specificity) is obtained as a consequence.
2. The required false positive rate is set, and the associated sensitivity is obtained as a consequence.
3. Equal values of sensitivity and specificity are set (when sensitivity and specificity are both important at the same level).

Set the threshold according to method 3, but critically think if this choice is correct for the COVID-19 test.

Final activity

Repeat the exercise for Test1.

Which of the two tests is better? or are they equivalent?