Evaluation of test performances: ROC analyses, etc.

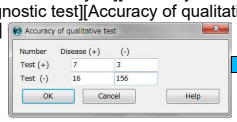
- To develop new method to detect diseases, what you need are:
  - Calculating sensitivity and specificity
    - Data: (if originally category data) Positive/Negative by that test, Truly disease/healthy by the gold standard
      - Sensitivity = Positive in Disease / All of Disease
      - Specificity = Negative in Healthy / All of Healthy
    - Data: (if originally continuous data) Values by that test, Truly disease/healthy by the gold standard
      - ROC analysis: by changing threshold value of positive/negative, seeking the best threshold as closest point to the upper left point where "sensitivity=1" and "1-specificity=0".
  - Compare several methods by ROC analysis
    - The method to achieve highest area under the curve (AUC) is judged as the one with best performance.
    - Actual determination of method may also consider cost, feasibility, etc.

Example1. Performance of malaria RDT for low parasite density

- Several RDTs (Rapid Diagnostic Tests) for malaria, originally developed to distinguish malaria patients from other fever patients
  - Patients with fever must have malaria parasites with high densi in their blood
    → High specificity and moderate sensitivity
- Is it also useful in active case detection study in low parasite density (less than 100 parasites / L) ?
- Pan-R malaria's results for P.vivax in Solomon Is. shown below [Statistical analysis][Accuracy of diagnostic test][Accuracy of qualitative test]

	疾病	健康
陽性	a	b
陰性	c	d

- 感度 (sensitivity) = a/(a+c)　※ positive in disease と覚えるといい。
- 偽陰性率= c/(a+c) = 1 − 感度
- 特異度 (specificity) = d/(b+d)　※ negative in health と覚えるといい。
- 偽陽性率= b/(b+d) = 1 − 特異度
- 陽性反応的中度 (positive predictive value) = a/(a+b)
- 陰性反応的中度 (negative predictive value) = d/(c+d)
- 陽性尤度比= (a/(a+c))/(b/(b+d)) =感度/(1 − 特異度)
- 陰性尤度比= (d/(b+d))/(c/(a+c)) =特異度/(1 − 感度)



Output Window

```
> .Table
          Disease positive Disease negative
Test positive                7              3
Test negative               16            156

> summary.test
                                    Estimation Lower 95%CI Upper 95%CI
Sensitivity                              0.304      0.132      0.529
Specificity                              0.981      0.946      0.996
Positive predictive value                0.700      0.348      0.933
Negative predictive value                0.907      0.853      0.946
Diagnostic accuracy                      0.896      0.842      0.936
Likelihood ratio of a positive test     16.130      4.485     58.008
Likelihood ratio of a negative test      0.709      0.541      0.930
```

Example 2. Determination of numerical criterion for diagnosis

- By the depression score based on the questionnaire, screen major depression.
- Requirement: Both patients who were clinically diagnosed as depression and not depression. The depression scores for them. (2nd line show the score, 3rd line is clinical diagnosis)

1	2	3	4	5	6	7	8	9	10
20	13	19	21	22	28	11	25	16	19
dep	norm	norm	norm	dep	dep	norm	norm	norm	norm

- If we set criterion as "more than 18 is depression", the cross table of diagnosis below. Sensitivity is 1, specificity is 3/7

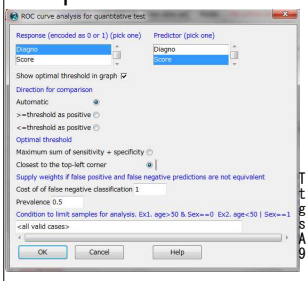
	Depression	Normal
Positive	3	4
Negative	0	3

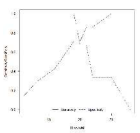
- By changing criteria, we can get the highest set of sens/spec

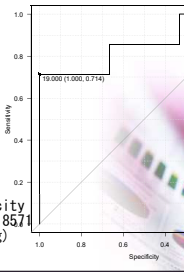
Example 2 (cont'd) ROC analysis of the depression score

- Enter the table from [File][New data set] as shown in the right screen cap.
- [Statistical analysis][Accuracy of diagnostic test][ROC …] and specify options like bottom-left screen cap.

	Score	Diagno
1	20	Dep
2	13	Norm
3	19	Norm
4	21	Norm
5	22	Dep
6	28	Dep
7	11	Norm
8	25	Norm
9	16	Norm
10	19	Norm







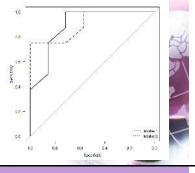
The criterion "more than 19 is depression" gives the best pairs of sensitivity and specificity
Area under the curve: 0.8571
95% CI: 0.6044-1 (DeLong)

Example 3. Compare several methods by ROC analysis

- The results of 2 different tests to evaluate the same thing may differ. We can compare them by AUC as the result of ROC analysis.

PID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pathology	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
Marker1	2.2	1.8	2.4	2.2	1.7	2.5	2.9	2.3	1.8	1.1	1.3	1.4	2.3	1.0	0.8
Marker2	3.5	2.8	3.9	3.4	1.8	3.0	3.1	2.0	2.1	0.9	2.7	0.9	2.0	0.5	0.4

- Note: The name of dataset must not be ROC1 nor ROC2. If you do so, those are overwritten during calculation to cause error.
- [Statistical analysis][Accuracy of diagnostic test] [Compare two ROC curves]
- Z = -0.0981, p-value = 0.9218
  AUC of roc1 AUC of roc2
  0.8928571 0.9017857



Agreements of 2 quantitative measurements

- Reliability of newly developed cheap or rapid measurement method has to be confirmed. For that purpose, agreement of the results (X) by new method obtained for the same subjects with the results obtained by the gold standard (Y)
- Method of checking
  - Paired t-test: cannot detect the interaction with absolute quantity
  - Correlation analysis and scatter gram to check the match with the line of x=y.
  - BA plot (Bland-Altman plot) is very famous. Since 1985, this is *de facto* standard.

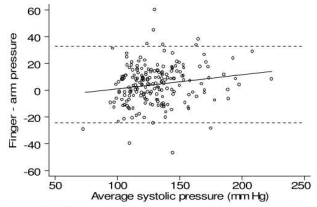


Figure 2. Difference against average of test and standard measurements, with 95% limits of agreement (broken lines) and regression line

Basically, make scattergram with (X-Y) as y-axis, (X+Y)/2 as x-axis.

How to draw BA-plot in R

- By manipulating 2 variables, it's possible to calculate the 2 new variables D (as X-Y) and M (as (X+Y)/2). Draw scattergram of M as x-axis and D as y-axis.
- Using MethComp package, BA.plot()
- library(MethComp)
  data(ox)
  BA.plot(ox)
- ox is the results of blood oxygen saturation of 61 children using gas (CO) measurements and pulse measurements