Advanced Statistics
DS2003 (BDS-4A)
Lecture 28

Instructor: Dr. Syed Mohammad Irteza

Assistant Professor, Department of Computer Science, FAST

02 June, 2022

Previous Lecture

- Example of Dr. House and the question, "its never really lupus"
  - Using our knowledge of Bayes' Theorem, we can construct a tree, and then find out the prob(lupus|+)
- Definitions:
  - True Positive
  - True Negative
  - False Positive
  - False Negative
  - Sensitivity and Specificity
- Spam Emails
  - Deciding a threshold

Recap

- Sensitivity
- Specificity
- True Positive
- True Negative
- False Positive
- False Negative

Hypothetical Example 1 - Screening Test A

- 100 people are tested for the disease. 15 people have the disease; 85 people are not diseased. So, prevalence is 15%:

- Sensitivity is two-thirds, so the test is able to detect two-thirds of the people with the disease. The test misses one-third of the people who have the disease.

- The test has 53% specificity. In other words, out of 85 persons without the disease, 45 have true negative results while 40 individuals test positive for a disease that they do not have.

Hypothetical Example 1 - Screening Test A

- Sensitivity:
  A/(A + C) × 100
  10/15 × 100 = 67%

- Specificity:
  D/(D + B) × 100
  45/85 × 100 = 53%

- Positive Predictive Value:
  A/(A + B) × 100
  10/50 × 100 = 20%

	Truth		
	Disease (number)	Non Disease (number)	Total (number)
Positive (number)	10 **A** (True Positive)	40 **B** (False Positive)	50 $T_{Test\ Positive}$
Negative (number)	5 **C** (False Negative)	45 **D** (True Negative)	50 $T_{Test\ Negative}$
	15 $T_{Disease}$	85 $T_{Non\ Disease}$	100 Total

Test Result

Hypothetical Example 2 - Increased Prevalence, Same Test

- 100 people are tested for the disease. 30 people have the disease; 70 people are not diseased. *So, prevalence is now 30%*:

- Sensitivity is two-thirds, so the test is able to detect two-thirds of the people with the disease. The test misses one-third of the people who have the disease.

- The test has 53% specificity. In other words, out of *70 persons* without the disease, *37* have true negative results while *33* individuals test positive for a disease that they do not have.

Hypothetical Example 2 - Increased Prevalence, Same Test

- Sensitivity:
  A/(A + C) × 100
  20/30 × 100 = 67%

- Specificity:
  D/(D + B) × 100
  37/70 × 100 = 53%

- Positive Predictive Value:
  A/(A + B) × 100
  20/53 × 100 = 38%

Truth

Test Result

	Disease (number)	Non Disease (number)	Total (number)
Positive (number)	20 **A** (True Positive)	33 **B** (False Positive)	53 $T_{Test\ Positive}$
Negative (number)	10 **C** (False Negative)	37 **D** (True Negative)	47 $T_{Test\ Negative}$
	30 $T_{Disease}$	70 $T_{Non\ Disease}$	100 **Total**

Useful Links & Resources

- **Reference**:
  - https://online.stat.psu.edu/stat507/lesson/10/10.3