# Logical Reasoning
## (part 2)

*Davide Balzarotti*
*davide@iseclab.org*


# Common Pitfalls

# Base Rate Fallacy

- A tool to detect attacks have
  - False positive rate of 0.1%
    (i.e., the probability that it flags a normal event as malicious is 1 in 1.000)
  - Detection rate of 99.9%
    (i.e., the probability that an attack is detected is 999 in 1.000)

# Base Rate Fallacy

- A tool to detect attacks have
  - False positive rate of  0.1%
    (i.e., the probability that it flags a normal event as malicious is 1 in 1.000)
  - Detection rate of 99.9%
    (i.e., the probability that an attack is detected is 999 in 1.000)

# Base Rate Fallacy

- A tool to detect attacks have
  - False positive rate of 0.1%
    (i.e., the probability that it flags a normal event as malicious is 1 in 1.000)
  - Detection rate of 99.9%
    (i.e., the probability that an attack is detected is 999 in 1.000)
- When an alert is generated, what is the probability that it was an attack?

# Base Rate Fallacy

- A tool to detect attacks have
  - False positive rate of 0.1%
    (i.e., the probability that it flags a normal event as malicious is 1 in 1.000)
  - Detection rate of 99.9%
    (i.e., the probability that an attack is detected is 999 in 1.000)
- When an alert is generated, what is the probability that it was an attack?

99.9 %

# Base Rate Fallacy

- A tool to detect attacks have
  - False positive rate of 0.1%
    (i.e., the probability that it flags a normal event as malicious is 1 in 1.000)
  - Detection rate of 99.9%
    (i.e., the probability that an attack is detected is 999 in 1.000)
- When an alert is generated, what is the probability that it was an attack?

~~99.9 %~~

Confusing `P(x|y)` with `P(y|x)` is a common mistake

- Especially severe when the probability of certain events is very low

# Base Rate Fallacy

$$P(A|F) = \frac{P(F|A) \cdot P(A)}{P(F)}$$

$$P(A|F) = \frac{P(F|A) \cdot P(A)}{P(F|A) \cdot P(A) + P(F|\neg A) \cdot P(\neg A)}$$

- The result depends on the base rate ( `P(A)` and `P(¬A)` )

# Base Rate Fallacy

$$P(A|F) = \frac{P(F|A) \cdot P(A)}{P(F)}$$

$$P(A|F) = \frac{P(F|A) \cdot P(A)}{P(F|A) \cdot P(A) + P(F|\neg A) \cdot P(\neg A)}$$

- The result depends on the base rate ( `P(A)` and `P(¬A)` )
- If we estimate that only 1 event in 100K is an attack,
  the previous formula tells us that **`P(A|F) = ~ 1%`**

# Multiple Testing Fallacy

- In a murder investigation, the killer blood sample is collected on the crime scene
- The probability that two DNA match by chance is 1 in 10.000
- The police run the DNA against the database of 20K DNAs of the village citizens and a positive match is found!

# Multiple Testing Fallacy

- In a murder investigation, the killer blood sample is collected on the crime scene
- The probability that two DNA match by chance is 1 in 10.000
- The police run the DNA against the database of 20K DNAs of the village citizens and a positive match is found!
- What is the probability that the person is not guilty?

# Multiple Testing Fallacy

- In a murder investigation, the killer blood sample is collected on the crime scene
- The probability that two DNA match by chance is 1 in 10.000
- The police run the DNA against the database of 20K DNAs of the village citizens and a positive match is found!
- What is the probability that the person is not guilty?

1 in 10.000 ?? NO

Multiple testing fallacy may occur when an evidence is compared against a large database. The probability of a match by change in the entire database is:

$$1-\left(1-\frac{1}{10000}\right)^{20000} \approx 0.86$$

# *Reasonable Doubt*

- Our goal is to reconstruct the events on a computer system until we reach a reasonable certainty
- Unfortunately, the existence of an artifact rarely tells you *how* or *why* it is there
- So, how do you know what is “*reasonable*” ?
  - Did you list all possible scenarios?
  - Under which assumptions you reached your conclusion?
  - What is your threat model?
  - How expert / motivated / funded are all involved actors?

# The Research Corner

"*Overcoming Reasonable Doubt in Computer Forensic  Analsysis*"
Jim Garrett – *SANS Computer Forensic Technical Paper*

"*The base rate Fallacy and the Difficulty of Intrusion   Detection*"
S.Axelsson - ACM CCS 1999

Start from the Wikipedia page on *Reasoning*,
then check the *Logical Fallacy* page