Engineering Statistics
IES 302

Dr. Prapun Suksompong

prapun@siit.tu.ac.th

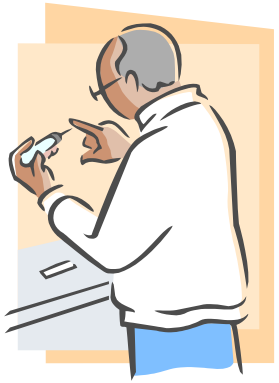
6.1 Conditional Probability

Suppose we have a diagnostic test for a particular disease which is 99% accurate. The test gives a positive result.

What is the probability that the person actually has the disease?

Disease Testing

- Suppose we have a diagnostic test for a particular **disease** which is 99% accurate.
- A person is picked at random and tested for the disease.
- The test gives a **positive result**.
- Q1: What is the probability that the person actually has the disease?
- Natural answer: 99% because the test gets it right 99% of the times.



99% accurate test?

- If you use this test on many persons **with** the disease, the test will indicate correctly that those persons have disease 99% of the time.
  - False negative rate = 1% = 0.01
- If you use this test on many persons **without** the disease, the test will indicate correctly that those persons do not have disease 99% of the time.
  - False positive rate = 1% = 0.01

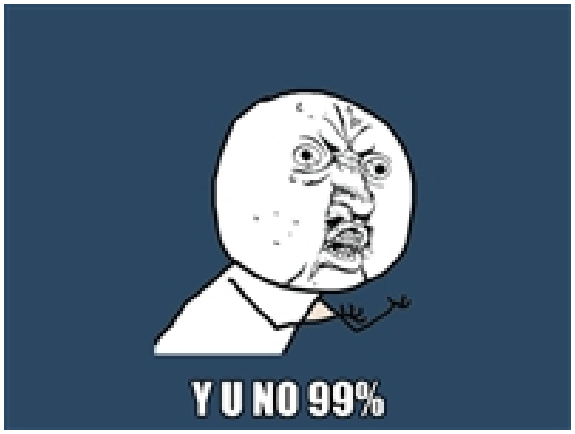
Disease Testing

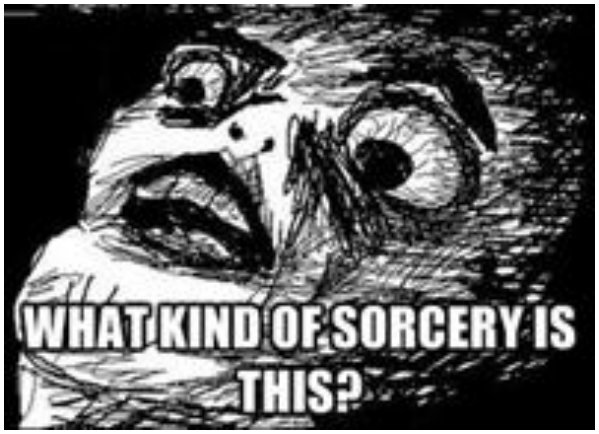
- Suppose we have a diagnostic test for a particular **disease** which is 99% accurate.
- A person is picked at random and tested for the disease.
- The test gives a **positive result**.
- Q1: What is the probability that the person actually has the disease?
- Natural answer: 99% because the test gets it right 99% of the times.
- Q2: Can the answer be 1% or 2%?
- Q3: Can the answer be 50%?

A1:

Q1: What is the probability that the person actually has the disease?

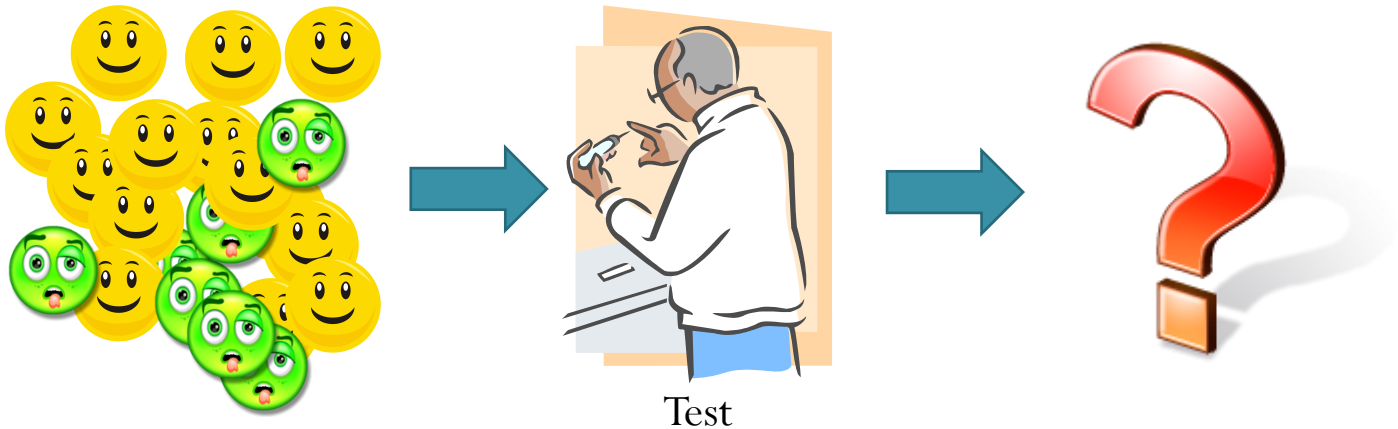
The answer actually depends on how **common** or how **rare** the disease is!





Why?

- Let's assume **rare disease**.
  - The disease affects about 1 person in 10,000.
- Try an experiment with 10^6 people.
- Approximately **100 people** will have the disease.
- What would the (99%-accurate) test say?



10^6 people

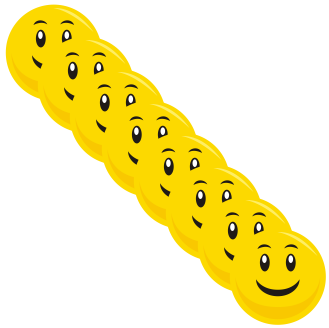
Results of the test

100 people w/ disease

approximately

99 of them will test positive

1 of them will test negative



999,900 people w/o disease

989,901 of them will test negative

9,999 of them will test positive

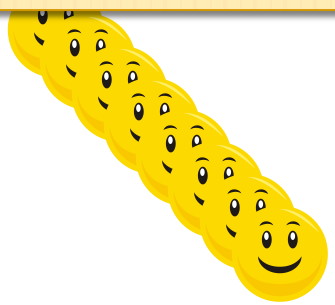
Results of the test

100 people w/ disease

99 of them will test positive

1 of them will test negative

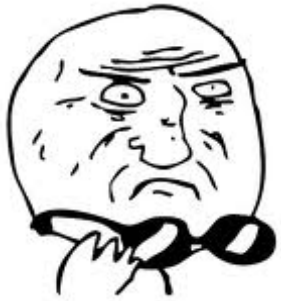
Of those who test positive, only $\frac{99}{99+9,999} \approx 1\%$ actually have the disease!



999,900 people w/o disease

989,901 of them will test negative

9,999 of them will test positive



Bayes' Theorem

Using the concept of **conditional probability** and **Bayes' Theorem**, you can show that

> the probability that a person will have the disease given that the test is positive

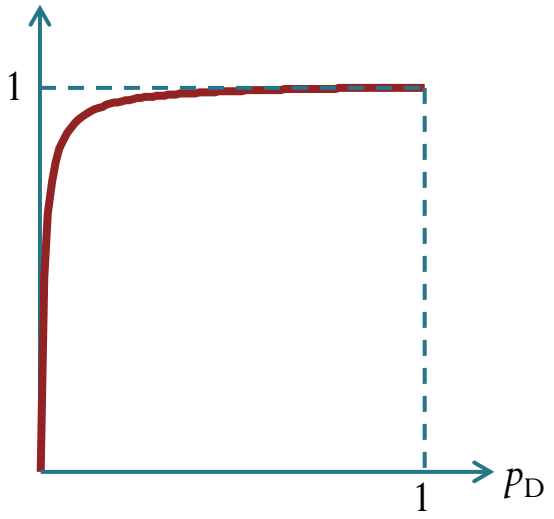
is given by

$$\frac{(1-p_{TE})\,p_D}{(1-p_{TE})\,p_D + p_{TE}(1-p_D)}$$

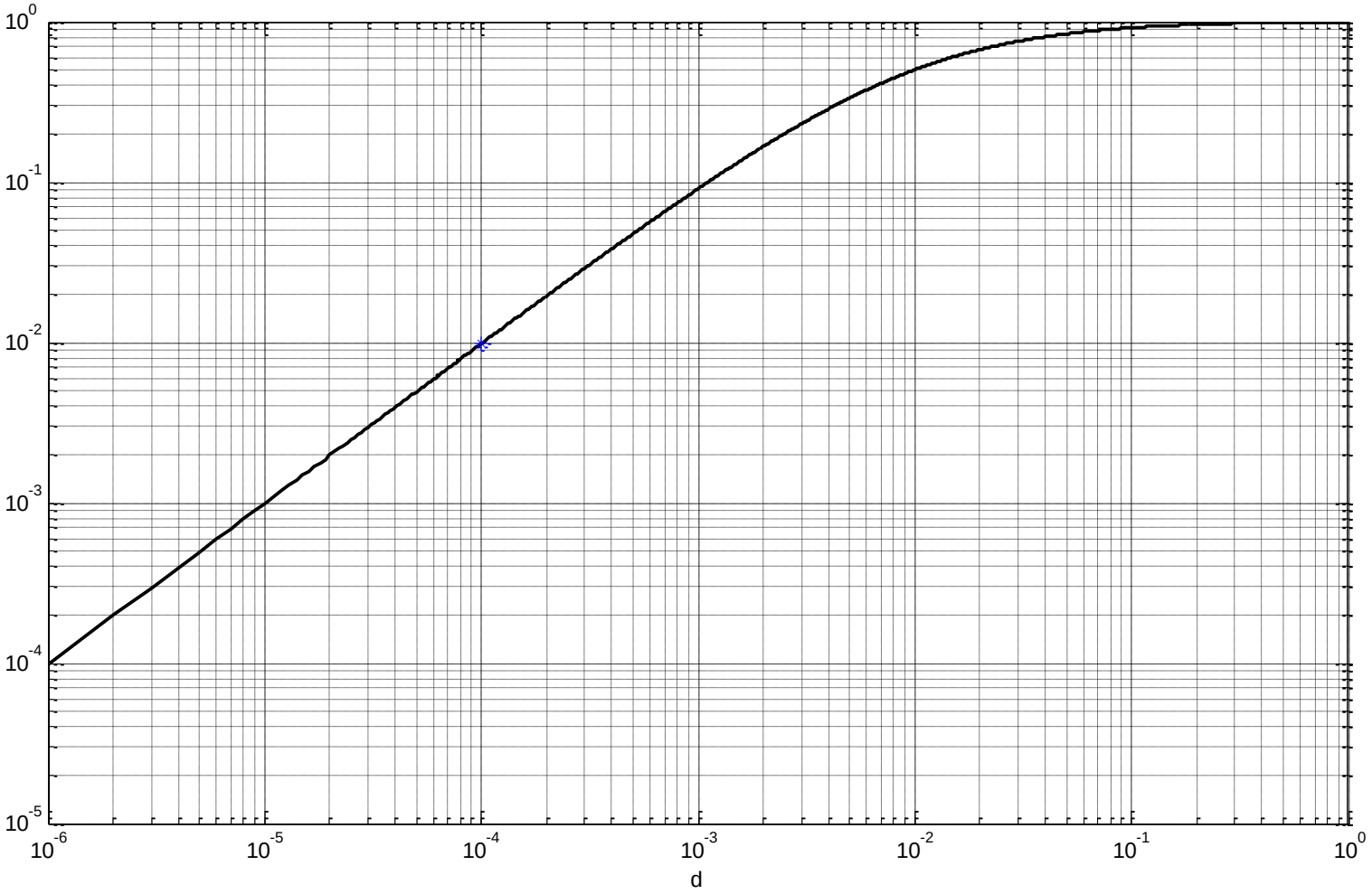
where

$p_D = 10^{-4}$

$p_{TE} = 1 - 0.99 = 0.01$



In log scale…

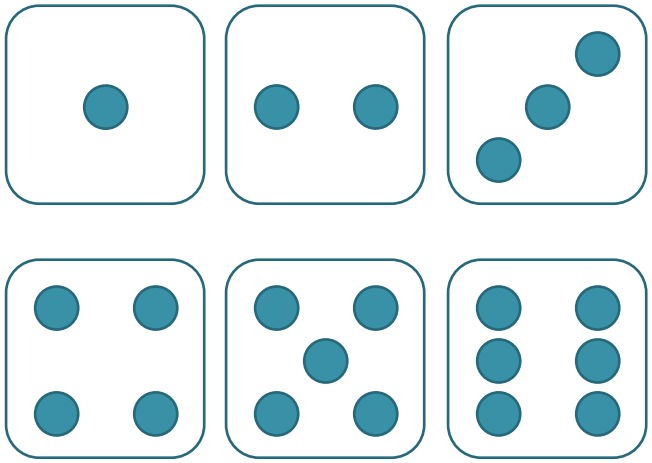


Wrap-up

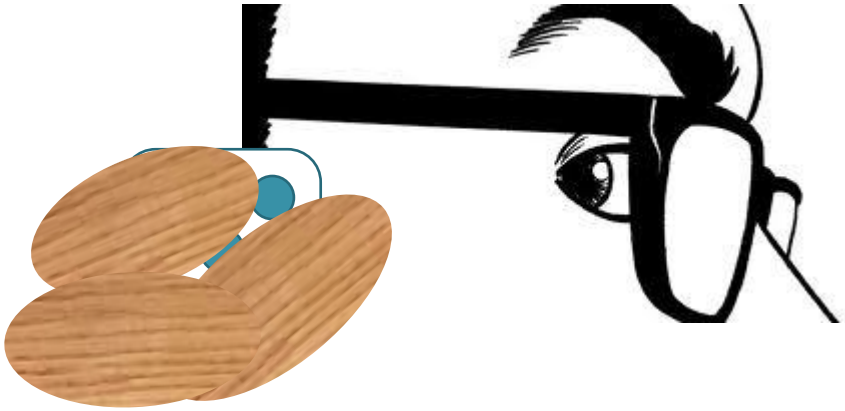
- Q1: What is the probability that the person actually has the disease?
- A1: The answer actually depends on how common or how rare the disease is! (The answer depends on the value of d.)
- Q2: Can the answer be 1% or 2%?
- A2: Yes.
- Q3: Can the answer be 50%?
- A3: Yes.

Example

- Roll a fair dice

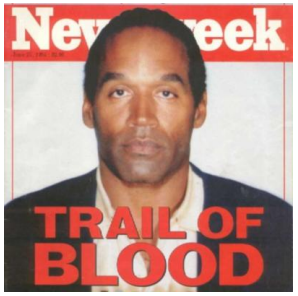


- Sneak peek:



Prosecutor's fallacy

- Murder case
  - "one of the biggest media events of 1994–95"
  - "the most publicized criminal trial in American history"
- O. J. Simpson
  - At the time a well-known celebrity famous both as a **TV actor** and as a retired **professional football star**.

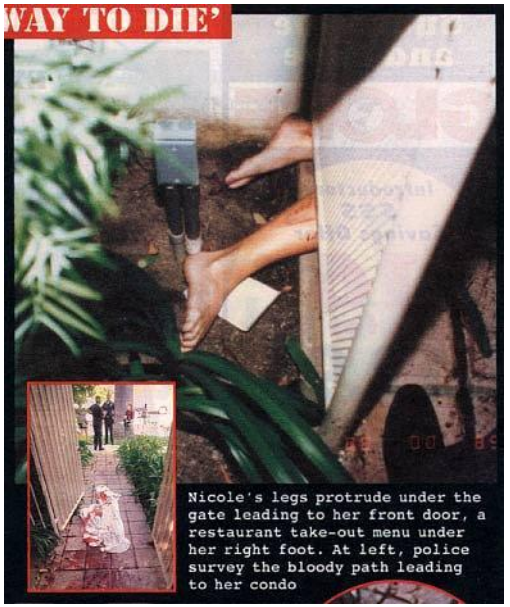


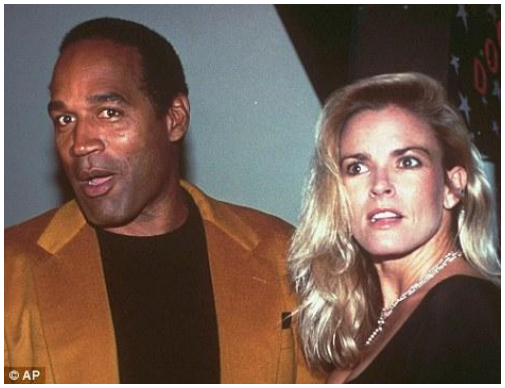
- Defense lawyer: Alan Dershowitz
  - Renowned attorney and **Harvard Law School professor**

[Mlodinow, 2008, p. 119-121],[Tijms, 1007, Ex 8.7]

The murder of Nicole

- **Nicole** Brown was murdered at her home in Los Angeles on the night of June 12, 1994.
  - So was her friend Ronald Goldman.
- The **prime suspect** was her (ex-) **husband** O.J. Simpson.
  - (They were divorced in 1992.)





Prosecutor* = a government official who conducts criminal prosecutions on behalf of the state

Prosecutors' argument

- Prosecutors* spent the first ten days of the trial entering **evidence** of Simpson's history of **physically abusing** her and claimed that this alone was a good reason to suspect him of her murder.
- As they put it,
  "a slap is a prelude to homicide."

prosecution = the lawyers acting for the state to put the case against the defendant
batter = strike violently and repeatedly

Counterargument

- The **defense attorneys** argued that the prosecution* had spent two weeks trying to **mislead** the jury and that the **evidence** that O. J. had battered Nicole on previous occasions **meant nothing**.
- **Dershowitz's reasoning**:
  - 4 million women are battered annually by husbands and boyfriends in the US.
  - In 1992, a total of 1,432, or 1 in 2,500, were killed by their (ex)husbands or boyfriends.
  - Therefore, few men who slap or beat their domestic partners go on to murder them.
- True? Yes.
- Convincing? Yes.
- Q: Was the fact that O.J. Simpson had previously physically abused his wife irrelevant to the case?

The verdict:

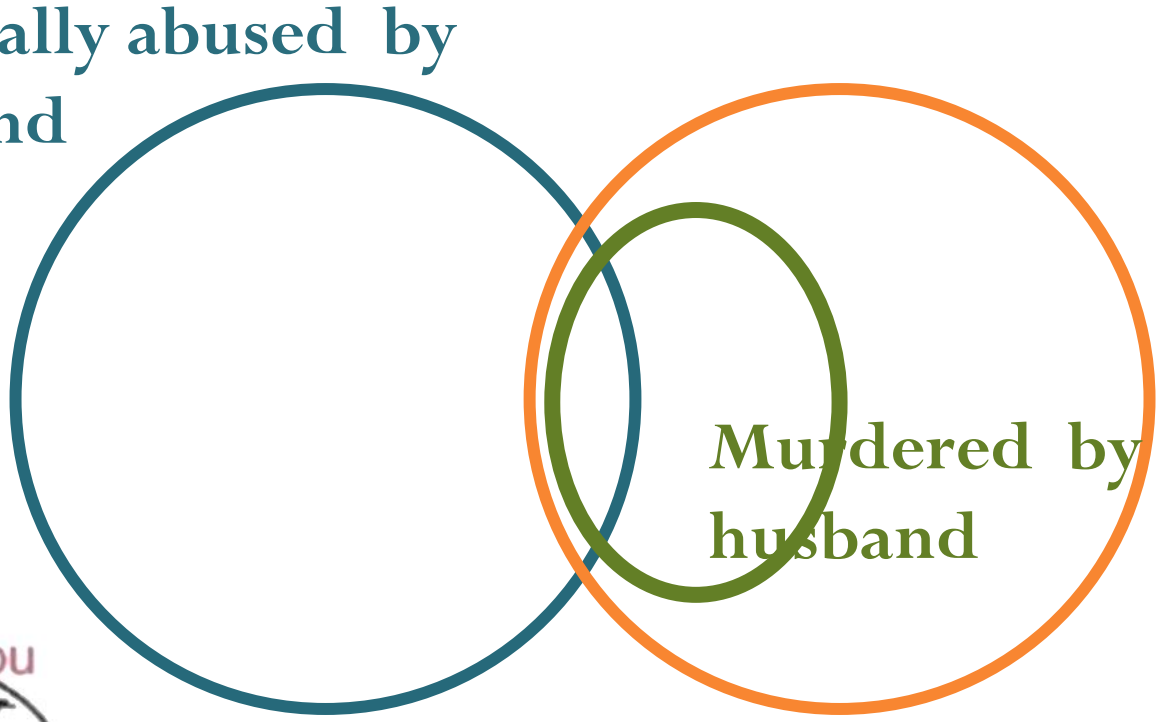
Not guilty for the two murders!

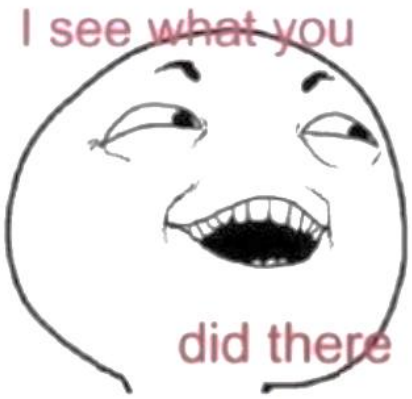
The verdict was seen live on TV by more than half of the U.S. population, making it one of the most watched events in American TV history.

Another number...

- It is important to make use of the crucial fact that Nicole Brown was murdered.
- The relevant number is not the probability that a man who batters his wife will go on to kill her (1 in 2,500) but rather the probability that a battered wife who was murdered was murdered by her abuser.
- According to the Uniform Crime Reports for the United States and Its Possessions in 1993, the probability Dershowitz (or the prosecution) should have reported was this one: of all the battered women murdered in the United States in 1993, some 90 percent were killed by their abuser.
- That statistic was **not mentioned at the trial**.

A Simplified Diagram





The Whole Truth …

- Dershowitz may have felt justified in **misleading** the jury because, in his words, "the courtroom oath—'to tell the truth, the ***whole truth*** and nothing but the truth'—is applicable only to witnesses.

- Defense attorneys, prosecutors, and judges don't take this oath . . . indeed, it is fair to say the American justice system is built on a foundation of not telling the whole truth."

Ex. Fair results from a biased coin

- A biased coin can still be used for fair results by changing the game slightly.
- John von Neumann gave the following procedure:
  1. Toss the coin twice.
  2. If the results match, start over, forgetting both results.
  3. If the results differ, use the first result, forgetting the second.
- Key idea:
  - The probability of getting heads and then tails must be the same as the probability of getting tails and then heads,
    - Assumptions: the coin is not changing its bias between flips and the two flips are independent.
  - By excluding the events of two heads and two tails by repeating the procedure, the coin flipper is left with the only two remaining outcomes having equivalent probability.
- This procedure only works if the tosses are paired properly; if part of a pair is reused in another pair, the fairness may be ruined.