Probabilities

Elias Zintzaras, M.Sc., Ph.D.

Professor in Biomathematics-Biometry
Department of Biomathematics
School of Medicine
University of Thessaly

Institute for Clinical Research and Health Policy Studies
Tufts University School of Medicine
Boston, MA, USA

Theodoros Mprotsis, MSc, PhD
Teacher & Research Fellow
(http://biomath.med.uth.gr)
University of Thessaly
Email: tmprotsis@uth.gr

The concept of probability

There is no complete and universally accepted explanation of the concept of probability (P), but in daily life, including medicine, we frequently use the term.

Example

- The result of antibiotic treatment is that the infection is either cured or not cured within 5 days
- A pathologist observed that 3 out of 4 patients were cured. Therefore, for this pathologist, the probability of a patient being cured with the antibiotic is 0.75 or 75%
- However, a larger sample size would provide a more accurate estimate of the treatment's effectiveness

Definition

Probability by definition is the frequency at which some event happens out of a greater number of outcomes

$$\boldsymbol{Probability = \frac{number\ of\ possible\ outcomes}{total\ number\ of\ outcomes}}$$

Example

- For example, if out of 100.000 newborns, 51.000 were girls, the probability that a newborn is a girl is represented by the fraction, $\frac{51.000}{100.000} = 0.51$, i.e. $\boldsymbol{P(girl) = 0.51}$

- Therefore, the probability that a newborn is a boy is 0.49, i.e., $\boldsymbol{P(boy) = 0.49}$

Example

- The probability of drawing a ♥ from a deck of playing cards is

$$P(heart) = \frac{13}{52} = \frac{1}{4}$$

- There are 52 possible events (the total number of cards in the deck), and 13 of them are considered successes, i.e., all the hearts

Properties

- Probability values range from **0 to 1**. A probability cannot be **negative**
- The probability of an **impossible event** (i.e., an event that cannot happen) **is 0**, while a probability of **1 means an event is certain**
- The greater the probability of an event, the more likely it is to happen
- An event with a probability of $\frac{1}{5}$ or 0.20 means that it has a 1 in 5 chance of happening

Definitions and Properties

- The probability of observing an event A is denoted by $\mathrm{P(A)}$, and it satisfies the condition:

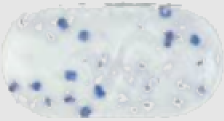
$$\mathbf{0 \le \boldsymbol{P(A)} \le 1}$$

- If events A, B, C, etc., are mutually exclusive (i.e., they cannot occur simultaneously, such as getting both heads and tails in a single coin toss), then the probability of any one of these events occurring is given by:

$$\boldsymbol{P(A\ or\ B\ or\ C\ \dots) = P(A) + P(B) + P(C) + \cdots}$$

Example

In pharma manufacturing, among 1.000.000 tablets in a batch, 50.000 are known to be flawed, perhaps containing specks of grease. Then, the probability of finding a randomly chosen tablet with specks is



$$P(specks) = 50.000/1.000.000 = 0.05\ (5\%)$$

Also, 30.000 have chipped edges and 40.000 are discolored. If these two defects are independent then,

$$P(chipped) = 30.000/1.000.000 = 0.03\ (3\%)$$

and

$$P(discolored) = 40.000/1.000.000 = 0.04\ (4\%)$$

Example

The probability of chosen an unacceptable tablet (specked, chipped, or discolored) at random is

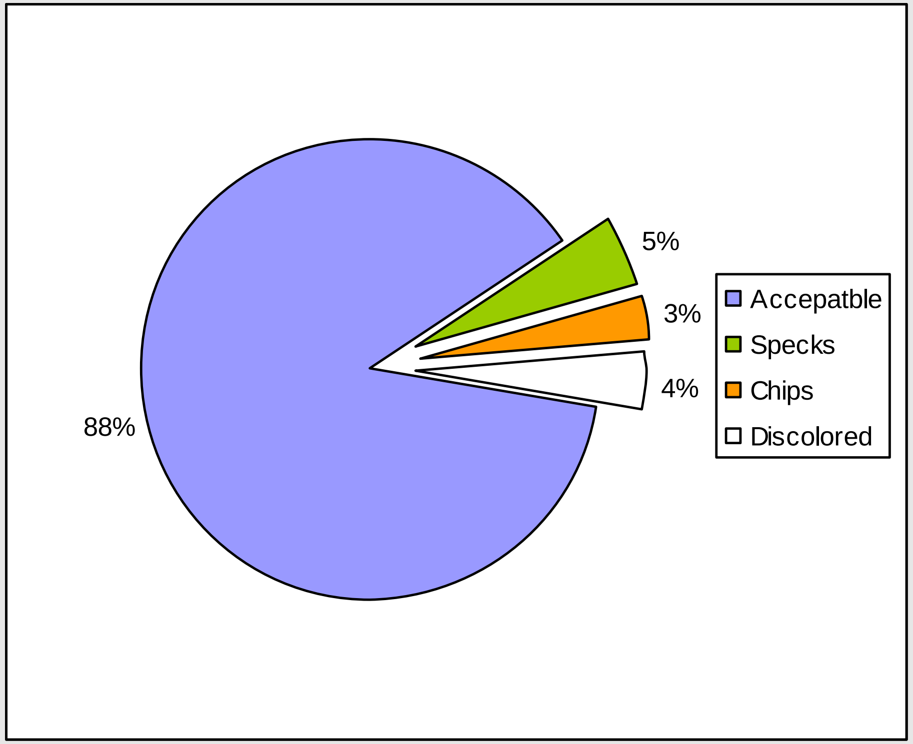
$$\boldsymbol{P(unacceptable\ tablet) =}$$

$$\boldsymbol{P(specked\ or\ chipped\ or\ discolored) =}$$

$$\boldsymbol{P(specked) + P(chipped) + P(discolored) =}$$

$$\mathbf{0.05 + 0.03 + 0.04 =}$$

$$\mathbf{0.12}$$



Mutually Exclusive Events

If the events A, B, C,… are mutually exclusive (i.e., they cannot occur simultaneously) and they cover all possible outcomes, then

$$\boldsymbol{P(A) + P(B) + P(C) + \cdots = 1}$$

In the example of the table attributes:

$$\boldsymbol{P(acceptable) + \ P(unaccetable) \ = \ 1}$$

Then, the probability of the chosen an acceptable tablet is

$$\boldsymbol{P(acceptable) \ = \ 1 \ - \ 0.12 \ = \ 0.88}$$

Mutually Exclusive Events

Table attributes	Acceptable	Specked	Chipped	Discolored	Total
Probability	0.88	0.05	0.03	0.04	1.00

Not Mutually Exclusive Events

If the events A and B are not mutually exclusive, then

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$P(\mathrm{A}\ or\, B)$ represents the probability that **either event A or event B occurs, or that both occur**

$P(A\ and\ B)$ represents the **probability that both events A and B occur at the same time**

If A and B are mutually exclusive (i.e., they cannot occur simultaneously), then

$$P(A\ and\ B) = 0$$

Example

In the example of tablet attributes, some chipped tablets may also be specked:

20.000 (2%) tablets are both chipped and specked

$P(specked\ or chipped)\ =$

$P(specked) + P(chipped) - P(specked\ and\ chipped) =$

$0.05 + 0.03 - 0.02 =$

$\mathbf{0.06}$

The probability of finding a specked or chipped tablet is $\mathbf{0.06}$ or $\mathbf{0.06 \cdot 1.000.000 = 60.000}$

Thus, it is expected $\mathbf{60.000}$ tablets to be specked or chipped

Conditional Probability

The conditional probability that an event A occurs given that another event B has already occurred, denoted as $P(\mathrm{A}|\mathrm{B})$, is given by the formula:

$$\boldsymbol{P(\mathrm{A}|\mathrm{B}) = \frac{P(A\ and\ \mathrm{B})}{P(\mathrm{B})}}$$

Then, the probability of both events (A and B) occurring is

$$\boldsymbol{P(A\ and\ B) = P(B) \cdot P(\mathrm{A}|\mathrm{B})}$$

Example: In the example of tablet attributes, the probability that a tablet will be specked given that the tablet is chipped is

$$\boldsymbol{P(specked|chipped) =}$$

$$\boldsymbol{P(specked\ and\ chipped)/P(specked)\ =}$$

$$\boldsymbol{\frac{0.02}{0.03} = \frac{2}{3} = 0.667 = 66.7\%}$$

This means that if a tablet is chipped, there is approximately 66.7% chance that it is also specked

Independent Events

The probability of two independent events (A and B) occurring together (i.e., the outcome of one event does not affect the other) is given by the product of their individual probabilities

$$P(A \; and \; B) = P(A) \cdot P(B)$$

Example

In the example of tablet attributes, the probability of selecting an acceptable tablet A followed by an unacceptable tablet B is

$$\boldsymbol{P(acceptable\ and\ unacceptable) =}$$
$$\boldsymbol{P(acceptable)\ \cdot\ P(unacceptable) =}$$
$$\boldsymbol{0.88 \cdot 0.12 = 0.106}$$

The probability of selecting two tablets, both of which are acceptable, is

$$\boldsymbol{P(acceptable\ and\ acceptable) =}$$
$$\boldsymbol{P(acceptable) \cdot P(acceptable) =}$$
$$\boldsymbol{0.88 \cdot 0.88 = 0.7744}$$

Probability Trees

A useful way of tackling many probability problems in medical decision is to draw a probability tree

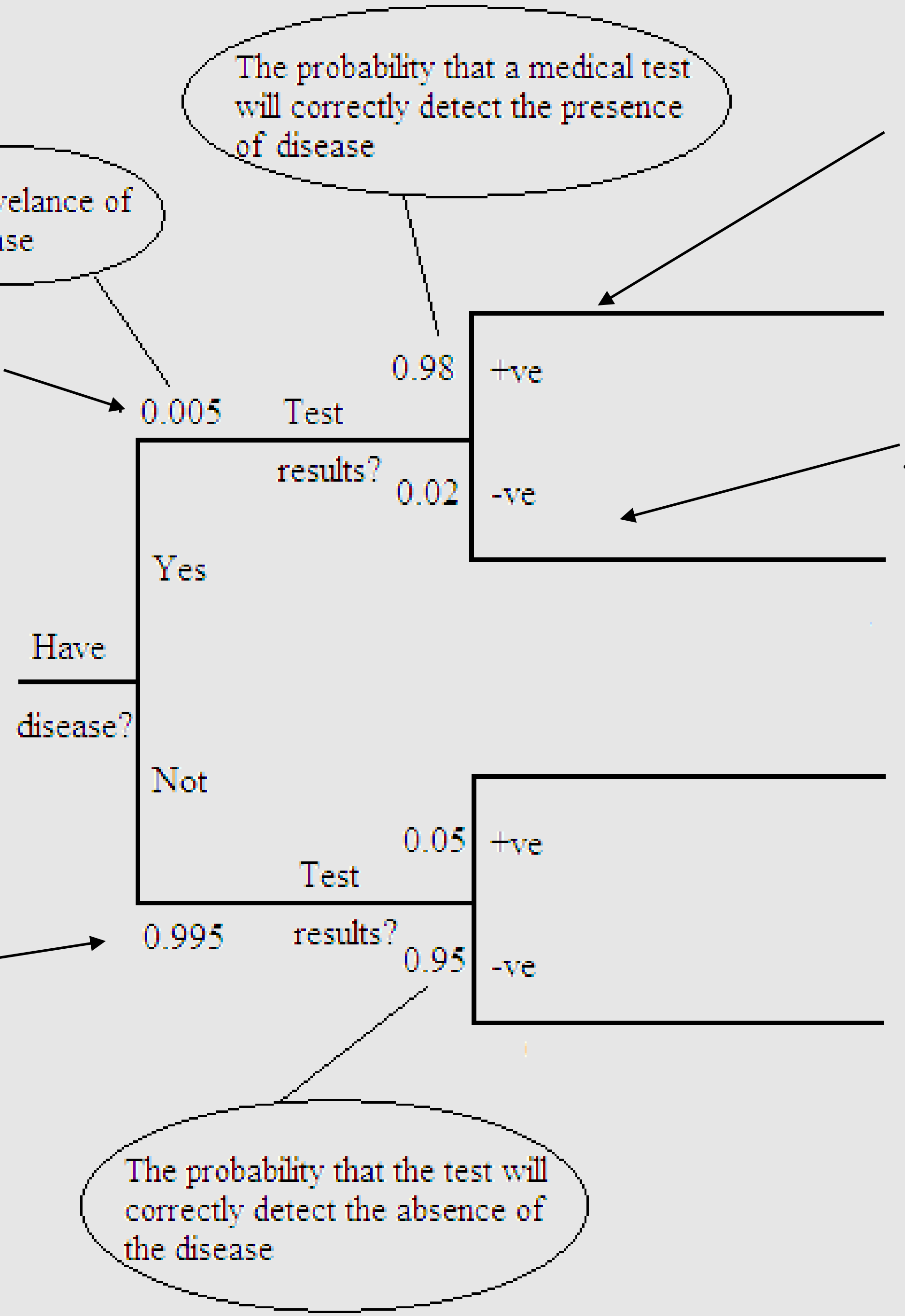
Example

- The probability that a medical test will correctly detect the **presence** of a certain disease is **98%**
- The probability that this test will correctly detect the **absence** of the disease is **95%**
- The disease is fairly rare, found in only **0.5%** of the population
- **If we have a positive test (meaning that the test says “yes, you got it”) what is the probability that we really have the disease?**

The disease is fairly rare, found in only 0.5% of the population

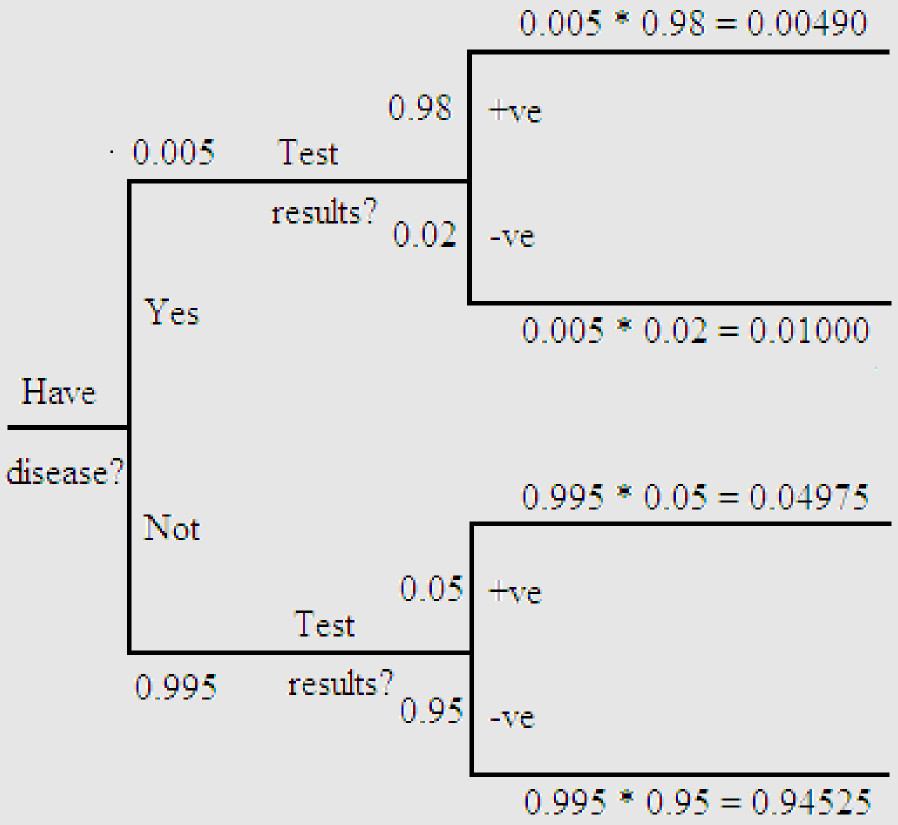
The first split in the tree represents the probability of the presence or absence of a disease in the population.

This means that 99.5% of the population **does not** have the disease.



If the patient has the disease, the probability that the test correctly detects it (true positive) is 0.98 (98%)

If the patient does not have the disease, the probability that the test correctly detects the absence of it (true negative) is 0.02 (2%)



The Probability of being diseased with +ve test is 0.00490

The Probability of being diseased with -ve test is 0.01000

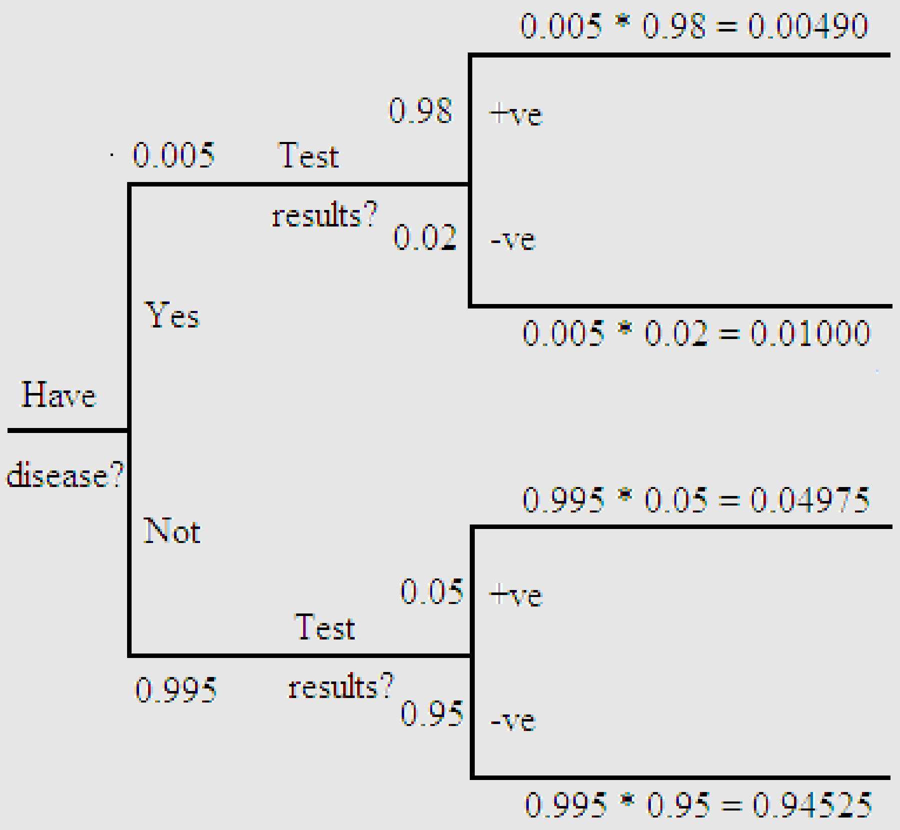
The Probability of being non-diseased with +ve test is 0.04975

The Probability of being non-diseased with -ve test is 0.94525

The probability of a positive test is 0.00490 + 0.04975 = 0.05465

The conditional probability of having the disease, given a positive test, is

P(Diseased | +ve test) = P(Diseased and +ve test) / P(+ve test) = 0.00490/0.05465 = 9%



Conclusion:

In general, a **9%** probability of actually having the disease after a positive test result is considered low.

It suggests that the test may not be sufficiently reliable, especially if the disease is serious.