Probability theory

Prof.dr.hab. Viorel Bostan

Technical University of Moldova

viorel.bostan@adm.utm.md

Lecture 3

Previous Lecture Outline

- Monty Hall Problem;
- Simulation of probability;
- Uniform distribution;
- Infinite sample space;
- Geometric distribution;
- Conditional Probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0;$$

- Product rule (why tree diagrams work).

$$P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B).$$

Family with children problem

Example

Consider a family with two children. **Given that one of the children is a boy, what is the probability that both children are boys?**

Suppose that the probability of a boy and a girl are the same (i.e. $1/2$).

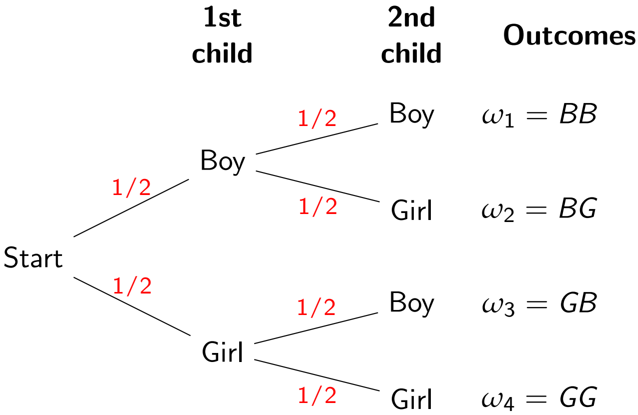
Your answer?

One way is to say that the other child is equally likely to be a boy or a girl, so the probability that both children are boys is $1/2$.

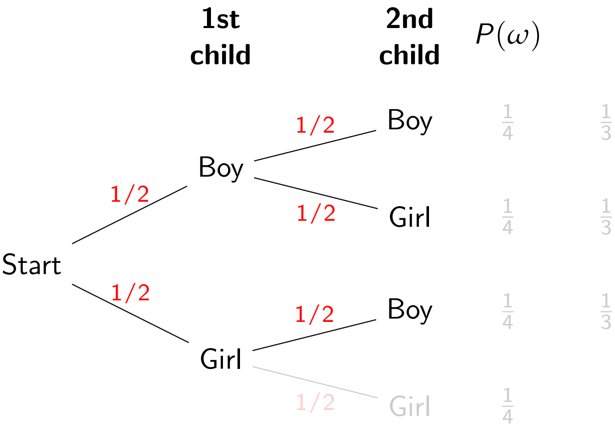
The “textbook” solution would be to draw the tree diagram and then form the conditional tree by deleting paths to leave only those paths that are consistent with the given information.

We will see that the probability of two boys given a boy in the family is not $1/2$, but rather... $\mathbf{1/3}$.

Family with children problem



Family with children problem



$A =$"one of the children is a boy" and $B =$"both children are boys".

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}$$

Another family problem

Example

Mr. Smith is the father of two.

We meet him walking along the street with a young boy whom he proudly introduces as his son.

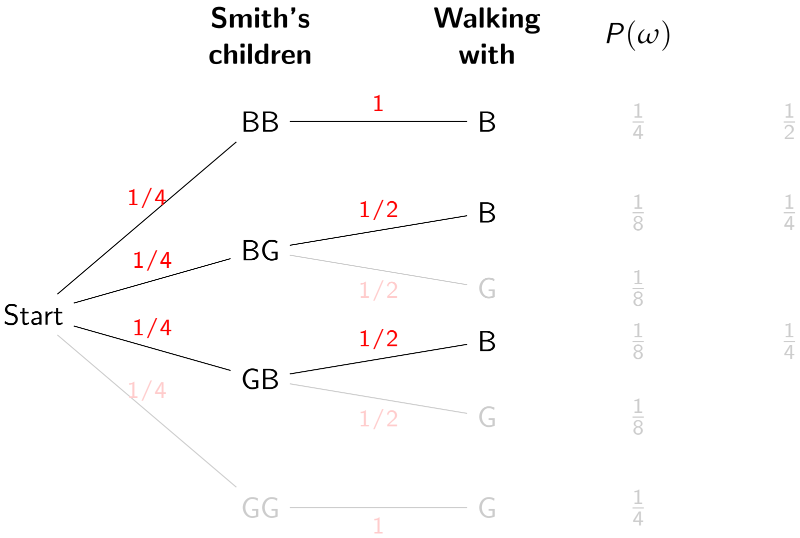
What is the probability that Mr. Smith's other child is also a boy?

What is your intuition whispering to you now?

As usual we have to make some additional assumptions.

For example, we will assume that if Mr. Smith has a boy and a girl, he is equally likely to choose either one to accompany him on his walk.

Consider the tree diagram to see that $1/2$ is the correct answer.


Smith's children
Walking with
$P(\omega)$
Start
1/4
1/4
1/4
1/4
BB
BG
GB
GG
1
1/2
1/2
1/2
1/2
1
B
B
G
B
G
G
$\frac{1}{4}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{8}$
$\frac{1}{4}$
$\frac{1}{2}$
$\frac{1}{4}$
$\frac{1}{4}$

Tiger–Rabbit problem

Example

A Tiger makes its lair in one of three caves:

The Tiger inhabits cave 1 with probability $1/2$, cave 2 with probability $1/4$, and cave 3 with probability $1/4$.

A rabbit subsequently moves into one of the two unoccupied caves, selected with equal probability.

With probability $1/3$, the rabbit leaves tracks at the entrance to its cave. (Tigers are much too clever to leave tracks.)

What is the probability that the Tiger lives in cave 3, given that there are no tracks in front of cave 2?

Start	T. choose cave	R. choose cave	R. leaves tracks?		Tiger in 3rd cave?
Start	Cave 1 (1/2)	Cave 2 (1/2)	Y (1/3)	1/12	
			N (2/3)	1/6	
		Cave 3 (1/2)	Y (1/3)	1/12	
			N (2/3)	1/6	
	Cave 2 (1/4)	Cave 1 (1/2)	Y (1/3)	1/24	
			N (2/3)	1/12	
		Cave 3 (1/2)	Y (1/3)	1/24	
			N (2/3)	1/12	
	Cave 3 (1/4)	Cave 1 (1/2)	Y (1/3)	1/24	Yes
			N (2/3)	1/12	Yes
		Cave 2 (1/2)	Y (1/3)	1/24	Yes
			N (2/3)	1/12	Yes

Tiger–Rabbit problem

Solution. Let B_3 be the event that the Tiger inhabits cave 3, and let T_2 be the event that there are tracks in front of cave 2. Thus, T_2^c denotes event that there no tracks in front of cave 2.

Taking data from the tree diagram, we can compute the desired probability as follows:

$$P(B_3 \mid T_2^c) = \frac{P(B_3 \cap T_2^c)}{P(T_2^c)}.$$

$$P(B_3 \cap T_2^c) = \frac{1}{24} + \frac{1}{12} + \frac{1}{12} = \frac{5}{24}.$$

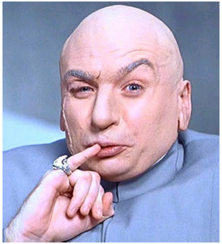
$$P(T_2^c) = 1 - P(T_2) = 1 - \left(\frac{1}{12} + \frac{1}{24}\right) = \frac{21}{24}$$

$$P(B_3 \mid T_2^c) = \frac{5/24}{21/24} = \boxed{\frac{5}{21}}.$$

Fiction problem

There are 3 prisoners in a maximum security prison for fictional villains:

- the Dark Lord Sauron,
- Doctor Evil,
- the Wolf that swallowed the little Red Hat and her grandma.


The parole board has declared that it will release 2 of the 3, chosen uniformly at random, but has not yet released their names.

Naturally, Sauron figures that he will be released to his home in Mordor, where the shadows lie, with probability 2/3.

Fiction problem

A corrupt guard offers to tell Sauron the name of one of the other prisoners who will be released (either Doctor Evil or Wolf).

However, Sauron declines this offer.

He reasons that if the guard says, for example, "Wolf will be released", then his own probability of release will drop down to 1/2.

This is because, he will then know that either he or Doctor Evil will also be released, and these two events are equally likely.

Is the Dark Lord Sauron reasoning correctly?

Define the events S, W, and GWR as follows:
$GWR =$ "Guard says Wolf is released";
$W =$ "Wolf is released";
$S =$"Sauron is released".

		Released		**Guard says**	$P(\omega)$	$\boldsymbol{GWR}$	$\boldsymbol{W}$	$\boldsymbol{S}$
Start	1/3	W,S	1	W	$1/3$	X	X	X
	1/3	W,D	1/2	W	$1/6$	X	X	
			1/2	D	$1/6$		X	
	1/3	D,S	1	D	$1/3$			X

Sauron is wrong!

Fiction Problem

Sauron's error is in failing to realize that event W (**Wolf will be released**) is different from event GWR (**guard says Wolf will be released**).

In particular, probability that Sauron is released, given that **Wolf is released**, is:

$$P(S \mid W) = \frac{P(S \cap W)}{P(W)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6} + \frac{1}{6}} = \frac{1}{2},$$

But the probability that Sauron is released given that **guard says Wolf is released** is

$$P(S \mid GWR) = \frac{P(S \cap GWR)}{P(GWR)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{6}} = \frac{2}{3}.$$

Sauron's probability of release is actually unchanged by guard's statement.

Law of Total Probability

Useful tool for computing probability by breaking into distinct cases.

Suppose we are interested in the probability of an event: $P(E)$.

Suppose also, that the random experiment can evolve in two different ways; that is, two different cases A and A^c are possible.

Assume that

- it is easy to find $P(A)$and $P(A^c)$;
- it easy to find $P(E \mid A)$ and $P(E \mid A^c)$.

Then, finding the probability of E is only two multiplications and one addition away.

Theorem (Law of Total Probability)

Let E and A be events, and $0 < P(A) < 1$. Then

$$P(E) = P(A)P(E \mid A) + P(A^c)P(E \mid A^c).$$

Law of Total Probability

Theorem (Law of Total Probability)

Let E and A be events, and $0 < P(A) < 1$. Then

$$P(E) = P(A)P(E \mid A) + P(A^c)P(E \mid A^c).$$

Proof.

$$\begin{aligned}
P(A)P(E \mid A) + P(A^c)P(E \mid A^c) &= \\
&= P(A)\frac{P(E \cap A)}{P(A)} + P(A^c)\frac{P(E \cap A^c)}{P(A^c)} \\
&= P(E \cap A) + P(E \cap A^c) \\
&= P((E \cap A) \cup (E \cap A^c)) \\
&= P(E \cap (A \cup A^c)) \\
&= P(E).
\end{aligned}$$

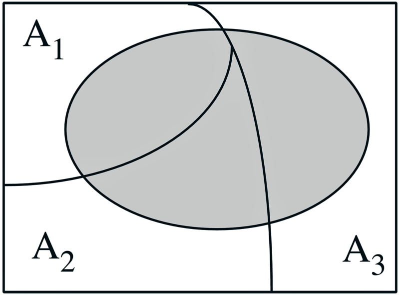
Law of Total Probability

Theorem

Let E be an event and let $A_1, A_2, \ldots, A_n$ be disjoint events such that $\cup_{i=1}^{n} A_i = \Omega$. Then,

$$P(E) = \sum_{i=1}^{n} P(E \mid A_i) P(A_i)$$

provided that $P(A_i) \neq 0$.



Rare disease problem

Example

There is a rare disease called **homeworkoholism**, which afflicts about one student in about 1000.

But it is worse than that!

At FAF the disease is making ravages afflicting **each third** student.

It is a compulsion to do as much homework as possible. It's horrible!

As victims are contaminated, they move downward spiral, and in four years they're awarded a degree in software engineering.

A doctor claim that he can diagnose homeworkoholism. Suppose you ask this doctor whether you have the disease.

Let $D =$ "you have the disease", and let $E =$ "the diagnosis is erroneous".

Rare disease problem

Example (Contd.)

If you have homeworkoholism, doctor says **yes** with probability 0.99:

$$P(\overline{E} \mid D) = 0.99$$
$$P(E \mid D) = 0.01$$

If you don't have it, doctor says **no** with probability 0.97.

$$P(\overline{E} \mid \overline{D}) = 0.97$$
$$P(E \mid \overline{D}) = 0.03$$

Compute $P(E)$, the probability that doctor makes a mistake.

Clearly at FAF: $P(D) = \frac{1}{3}$.

Obviously, D and $\overline{D}$ are disjoint and $D \cup \overline{D} = \Omega$.

Example (Contd.)

By **Total Probability Law**:

$$
\begin{aligned}
P(E) &= P(E \mid D)P(D) + P(E \mid \overline{D})P(\overline{D}) \\
&= 0.01 \cdot \frac{1}{3} + 0.03 \cdot \frac{2}{3} \\
&= 0.0233
\end{aligned}
$$

For a general student we have $P(D) = 0.001$. Thus,

$$P(E) = 0.01 \cdot 0.001 + 0.03 \cdot 0.999 = 0.02998 \approx 0.03$$

Previously, we have considered conditional probabilities of the following form:

Given the outcome of the 1st stage of a two-stage experiment, find the probability for an outcome at the 2nd stage.

But, often, there is a need to consider conditional probabilities of the following form:

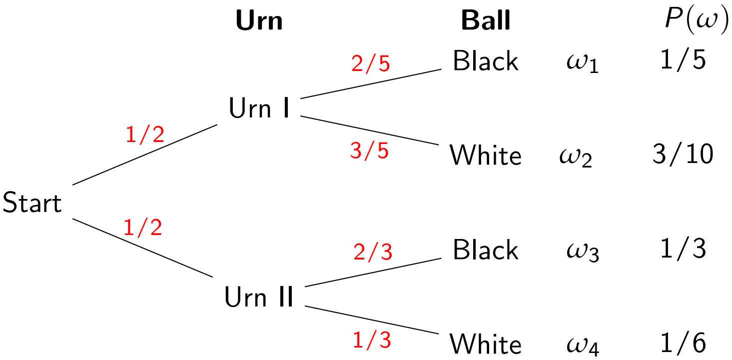
Given the outcome of the 2nd stage of a two-stage experiment, find the probability for an outcome at the 1st stage.

These probabilities are called **Bayes** probabilities or **a posteriori** probabilities. For example,

- **What is the probability that in the morning was cloudy if in the afternoon have rained?**
- **What is the probability that student did not prepared for the final exam if he/she have failed at the final exam?**

A posteriori probability

Consider the "urns and balls" problem from last lecture:



$$P(B) = m(\omega_1) + m(\omega_3) = \frac{1}{5} + \frac{1}{3} = \frac{8}{15},$$
$$P(W) = m(\omega_2) + m(\omega_4) = \frac{3}{10} + \frac{1}{6} = \frac{7}{15}.$$

A posteriori probability

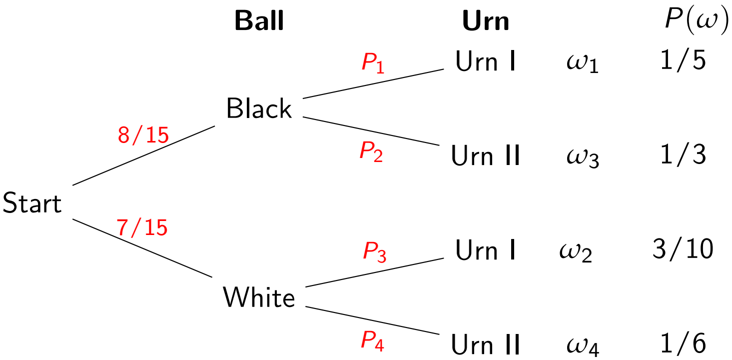
Suppose, we have drawn a black ball.

What is the probability that this ball was from first urn?

In other words, what is $P(I \mid B)$?

We can use the reverse tree approach, in which we will permute the stages: first will be color of the ball, and the second stage will be urn chosen.

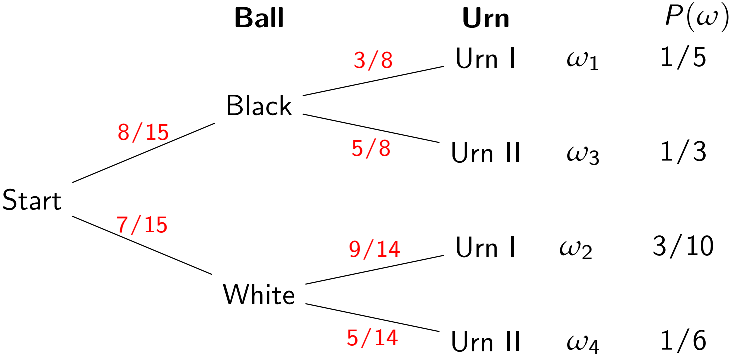
Reverse tree approach



How to compute P_1, P_2, P_3, P_4? We know that:

$$\frac{8}{15} * P_1 = \frac{1}{5},$$
$$P_1 = \frac{1/5}{8/15} = \frac{3}{8}.$$

Reverse tree approach



$$P(I \mid B) = \frac{3}{8}, \qquad P(II \mid B) = \frac{5}{8},$$
$$P(I \mid W) = \frac{9}{14}, \qquad P(II \mid W) = \frac{5}{14}.$$

Medical test problem

A doctor gives a patient a test for a A(H1N1) flu.

Before the results of the test, the only evidence the doctor has to go on is that 1 person in 1000 has this flu.

Experience has shown that, in 99 percent of the cases in which the flu virus is present, the test is positive; and in 95 percent of the cases in which it is not present, it is negative.

If the test turns out to be positive, what probability should the doctor assign to the event that the flu virus is present?

A priori probabilities are:

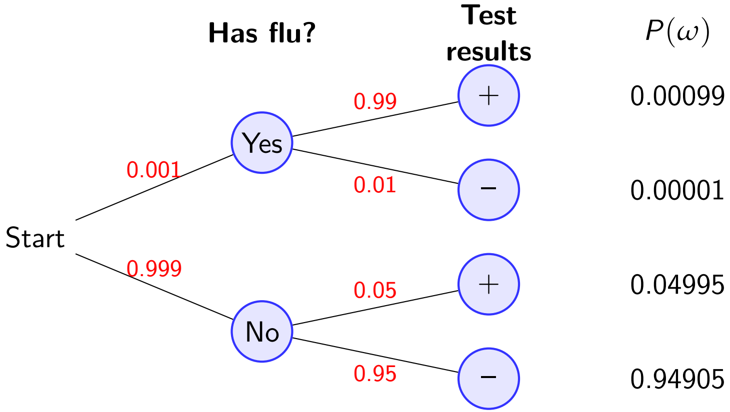
$$P(\mathit{flu}) = 0.001, \quad P(\mathit{not\ flu}) = 0.999.$$

We know also that

$$\begin{aligned} P(+ \mid \mathit{flu}) &= 0.99, \quad P(- \mid \mathit{flu}) = 0.01, \\ P(+ \mid \mathit{not\ flu}) &= 0.05, \quad P(- \mid \mathit{not\ flu}) = 0.95. \end{aligned}$$

Medical test problem

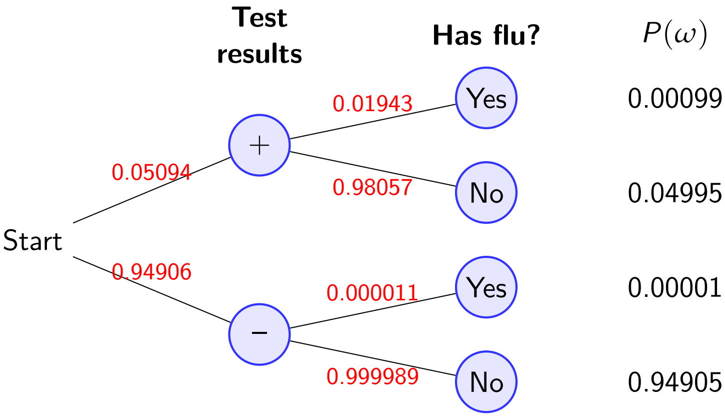
Using this data gives the result shown in the next tree diagram.



Need to compute a posteriori probability $P(\mathit{flu} \mid +)$.

Construct the reverse tree diagram.

Medical test problem. Reverse tree.



$$P(\mathit{flu} \mid +) = 0.019435$$

In other words, if the test results are positive the probability of having a flu virus increased from 0.001 to 0.019435: 20 times.

Medical test problem

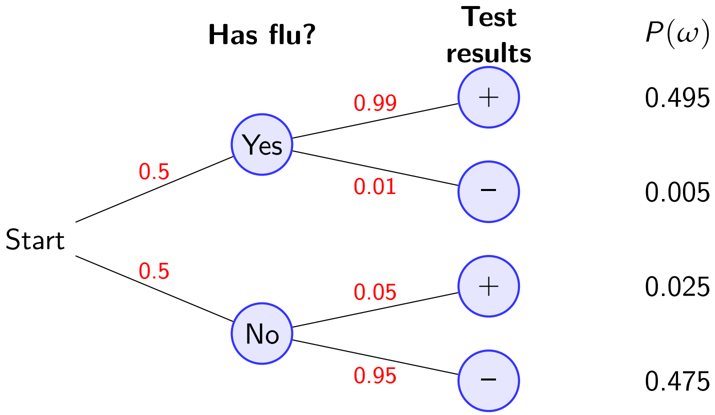
What if only people with symptoms are tested?

In that case, the evidence says that 50% will have flu.

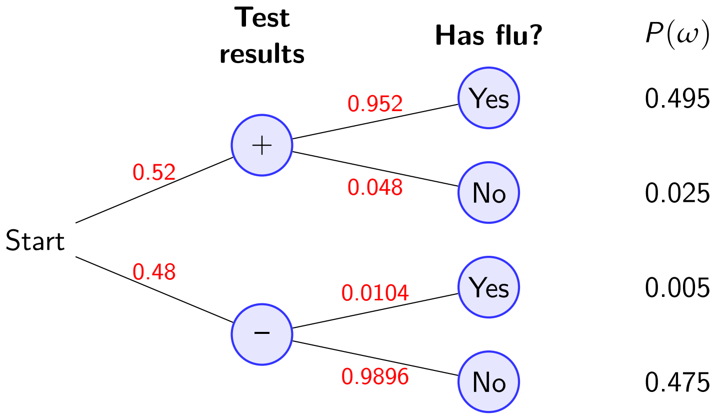
Construct tree and reverse tree diagrams.

Medical test problem

Using this data gives the result shown in the next tree diagram.



Medical test problem. Reverse tree.



$$P(flu \mid +) = 0.952$$

In other words, if the test results are positive the probability of having a flu virus $P(flu \mid +)$ increased from 0.5 to 0.952.

A posteriori Probabilities

Consider more general Bayes (a posteriori) probabilities.

Suppose we have a set of events $A_1, A_2, ..., A_m$ that are pairwise disjoint and such that the sample space satisfies the equation

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_m.$$

We call these events **hypotheses**.

We also have an event E that gives us some information about which hypothesis is correct and we call this event E **evidence**.

Before we receive the evidence, we have a set of prior probabilities $P(A_1)$, $P(A_2)$, $\ldots$, $P(A_m)$ for the hypotheses.

If we know the correct hypothesis, we know the probability for the evidence.

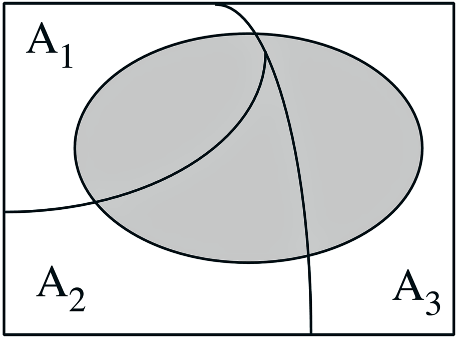
That is, we know $P(E|A_i)$ for all i.

A posteriori Probabilities

We want to find the probabilities for the hypotheses given the evidence.

That is, we want to find the conditional probabilities $P(A_i|E)$.

These probabilities are called **the a posteriori probabilities** (Bayes).



To find these probabilities, we write them in the form

$$P(A_i \mid E) = \frac{P(A_i \cap E)}{P(E)}. \tag{1}$$

We can calculate the numerator from given information by

$$P(A_i \cap E) = P(A_i)P(E|A_i). \tag{2}$$

Since one and only one of the events $A_1, A_2, ..., A_m$ can occur, we can write the probability of E (using formula (2) as

$$\begin{aligned} P(E) &= P(A_1 \cap E) + P(A_2 \cap E) + \ldots + P(A_m \cap E) \\ &= P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2) + \ldots + P(A_m)P(E \mid A_m) \\ &= \sum_{k=1}^{m} P(A_k)P(E \mid A_k). \end{aligned}$$

Substitute the last formula and (2) in (1) to get **Bayes** formula (on next slide).

Bayes Formula

Theorem (Bayes Formula)

$$P(A_i \mid E) = \frac{P(A_i)P(E|A_i)}{\sum_{k=1}^{m} P(A_k)P(E \mid A_k)},$$

or in extended form

$$P(A_i \mid E) = \frac{P(A_i)P(E|A_i)}{P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2) + \ldots + P(A_m)P(E \mid A_m)}.$$

Bayes Formula. Medical diagnosis example

Example

A doctor is trying to decide if a patient has one of three diseases D_1, D_2, or D_3 using two tests, each of which results in a positive $(+)$ or a negative $(-)$ outcome.

There are four possible test patterns: $++$, $+-$, $-+$, $--$.

National records have indicated that, for $10,000$ people having one of these three diseases, the distribution of diseases and test results are presented in the table below.

Disease	Nr. having this disease	The test results			
		$++$	$+-$	$-+$	$--$
D_1	3215	2110	301	704	100
D_2	2125	396	132	1187	410
D_3	4660	510	3658	73	509
Total	10000				

Bayes Formula. Medical diagnosis example

Disease	Nr. having this disease	The test results			
		$++$	$+-$	$-+$	$--$
D_1	3215	2110	301	704	100
D_2	2125	396	132	1187	410
D_3	4660	510	3658	73	509
Total	10000				

Estimate the prior probabilities for each of the diseases and, given a particular disease, the probability of a particular test outcome.

For example, the probability of disease D_1 may be estimated to be $P(D_1) = \frac{3215}{10{,}000} = 0.3215$.

The probability of the test result $+-$, given disease D_1, may be estimated to be $P(+- \mid D_1) = \frac{301}{3215} = 0.094$.

Bayes Formula. Medical diagnosis example

Use Bayes' formula to compute various posterior probabilities:

$$P(D_1 \mid ++) = \frac{P(D_1)P(++ \mid D_1)}{P(D_1)P(++ \mid D_1) + P(D_2)P(++ \mid D_2) + P(D_3)P(++ \mid D_3)}$$
$$= \frac{\frac{3215}{10000} \cdot \frac{2110}{3125}}{\frac{3215}{10000} \cdot \frac{2110}{3125} + \frac{2125}{10000} \cdot \frac{396}{2125} + \frac{4660}{10000} \cdot \frac{510}{4660}} = 0.7.$$

The results for this example are shown below:

	D_1	D_2	D_3
$++$	0.700	0.131	0.169
$+-$	0.075	0.033	0.892
$-+$	0.358	0.604	0.038
$--$	0.098	0.403	0.499

Lecture 3 Outline

- Conditional probability examples (Family with two boys, Tiger-rabbit, Fiction);
- Law of total probability;
- Example (Rare disease);
- A posteriori probabilities;
- Example (Medical test problem);
- Bayes formula;
- Example (Medical diagnosis problem).

Probability Joke

A student was completely unprepared for his final exam on probability. Since the exam was a True/False test, and recalling some things he heard in class, he decided to toss a coin for the answers.

The probability professor watched the student the entire two hours as he was tossing the coin... and writing the answer... tossing the coin...writing the answer.

At the end of the two hours, everyone else had left the classroom except for the one student.

The professor walks up to his desk and interrupts the student, saying:

"Listen, I have seen that you did not study for this probability test, you didn't even read the questions. If you are just tossing a coin for your answers, what is taking you so long to finish?

The student replies bitterly (as he is still tossing the coin): . . .

" Shhh! I am checking my answers!"