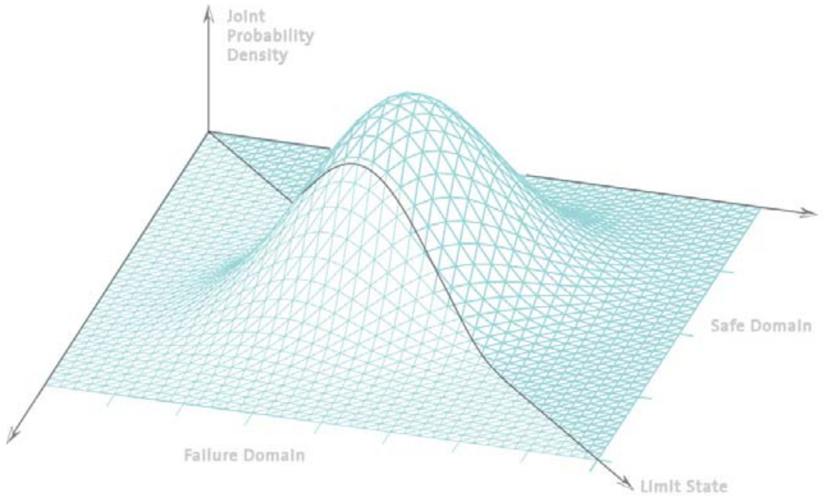


Exercises Tutorial 2

Statistics and Probability Theory

Prof. Dr. Michael Havbro Faber
Swiss Federal Institute of Technology Zurich
ETH Zurich

Introduction

Function and operation

In the school we have learned about "normal" functions:

		input	output	
	f	2	4	$f(x) = x^2$
Now				
	P	"1 comes out"	1/6	ideal dice

In the school we have learned about "normal" operations:

	$+, -, \times, \div$	for real numbers	:	2 + 4
Now				
	$\cap, \cup, \overline{}$	for set	:	$A \cap B$

Exercise 2.1 a)

Which of the following expressions are meaningful in the way they are written?

$$P\left[A \cup [B \cap C]\right]$$

$$P[A] + P[B]$$

$$P\left[\overline{A}\right] \cap P[B]$$

$$\overline{P[B]}$$

Answer 2.1 a)

Which of the following expressions are meaningful in the way they are written?

- ○ $P\left[A \cup [B \cap C]\right]$
- ○ $P[A] + P[B]$
- × $P\left[\overline{A}\right] \cap P[B]$ — P[ ] is a real number, while ∩ is an operation for set.
- × $\overline{P[B]}$ — P[ ] is a real number, while ¯ is an operation for set.

Probabilities cannot be separated and complementary events describe quantities and not the probability.

Exercise 2.1 b)

Assume A, B and C represent different events. Explain in words the meaning of the following expressions and what do they represent in mathematical terms (i.e. numbers, vectors, functions, sets,…)

$$A \cup B$$

$$\overline{B} \cap C$$

$$P[A]$$

$$P\Big[[A \cap B \cap C] \cup \big[\overline{A} \cap \overline{B} \cap \overline{C}\big]\Big]$$

$$\varnothing$$

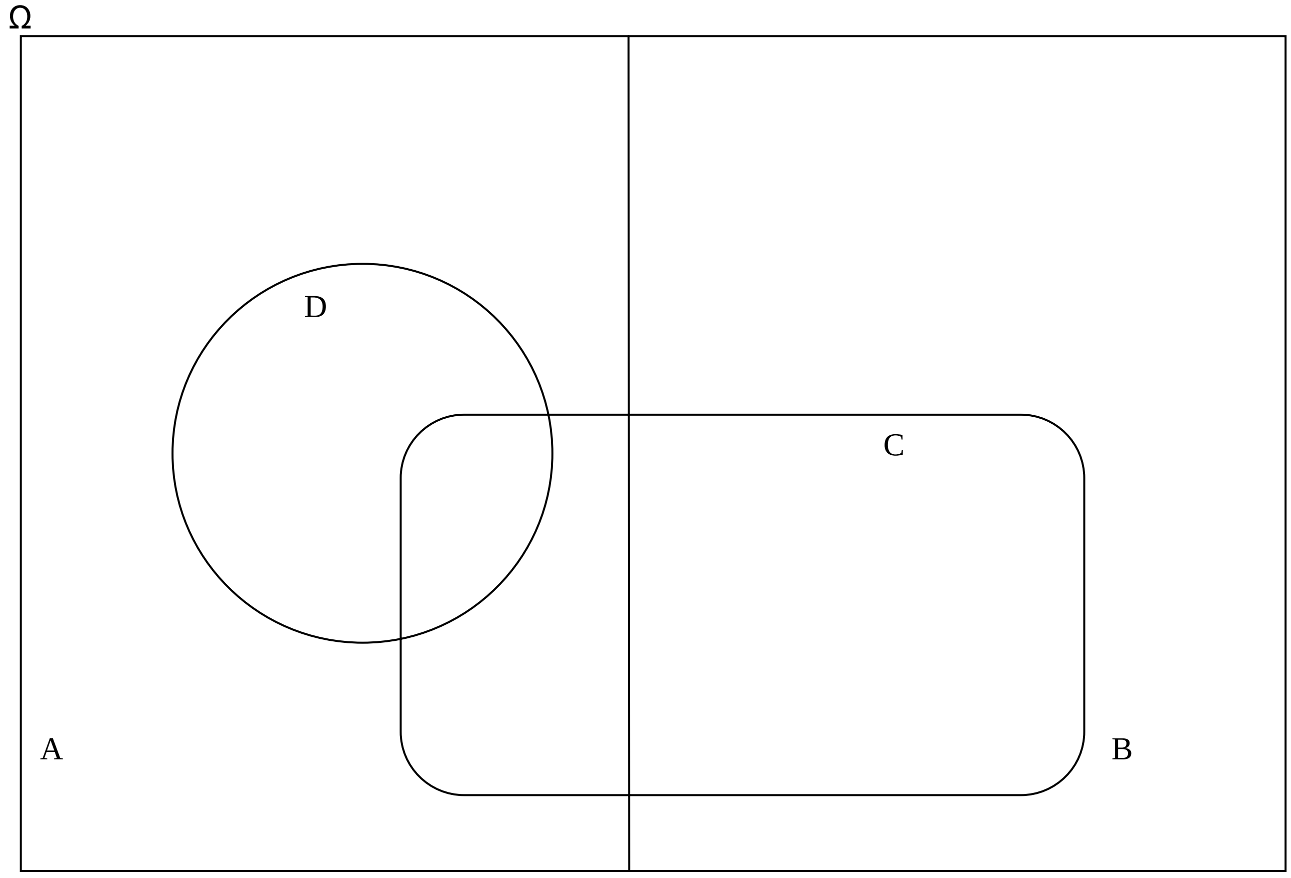
Exercise 2.1 c)

Using the diagram provided below show the following events.

$C \cap D$

$[D \setminus C] \cup [C \cap A]$

$B \cup D$



Exercise 2.2

We are throwing an ideal dice and considering the following events:

1. A: “An even number comes.”
   B: “A number dividable by 3 comes.”

2. A: “An even number comes.”
   B: “A prime number comes.”

Calculate the probability that the both events occur simultaneously for each case.

Exercise 2.2

We are throwing an ideal dice and considering the following events:

1. ***A*: "An even number comes."**
   ***B*: "A number dividable by 3 comes."**

Calculate the probability that both events occur simultaneously

$\Omega=\{1,2,3,4,5,6\}$

1. $A=\{2,4,6\}$, $B=\{3,6\}$, $A \cap B=\{6\}$

   $P(A)=1/2$, $P(B)=1/3$, $P(A \cap B)=1/6$

Exercise 2.2

We are throwing an ideal dice and considering the following events:

1. A: "An even number comes."
   B: "A number dividable by 3 comes."

2. **A: "An even number comes."**
   B: "A prime number comes."

Calculate the probability that both events occur simultaneously

$\Omega=\{1,2,3,4,5,6\}$

1. $A=\{2,4,6\}$, $B=\{3,6\}$, $A \cap B=\{6\}$

   $P(A)=1/2$, $P(B)=1/3$, $P(A \cap B)=1/6$

2. $A=\{2,4,6\}$, $B=\{2,3,5\}$, $A \cap B=\{2\}$

   $P(A)=1/2$, $P(B)=1/2$, $P(A \cap B)=1/6$

Definition of “independent”

$$P(A \cap B) = P(A) \cdot P(B)$$

but, what does it really mean??

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$$

What is the probability of B when you already know A?	=	What is the probability of B without any information?

The information on A cannot increase or decrease the probability of B.

More frankly, information on A is helpless, when you want to know about B in a probabilistic sense.

Why it is important to know about events being “independent” or not.

The story on an election some decades ago…

President election – two candidates A and B.

- TV station wanted to estimate the result of the election in advance.
- They asked those who voted which candidate they in fact voted, by telephone.

- Based on many many evidences, they concluded that: A won the election.
- However after counting the votes: B won the election.

What happened in reality,

- Those having a telephone were rich at that period.
- Rich people tended to vote candidate A.
- Poor people tended to vote candidate B.
- There were much more poor people than rich people.
- TV stations asked only by telephone!

What happened in reality,

- ➢Those having a telephone were rich at that period.
- ➢Rich people tended to vote candidate A.
- ➢Poor people tended to vote candidate B.
- ➢There were much more poor people than rich people.
- ➢TV stations asked only by telephone!

Let's assume that
T is the event that a person has a telephone.
V is the event representing which candidate he/she voted.

What they wanted to know is the probability of V : $P(V)$

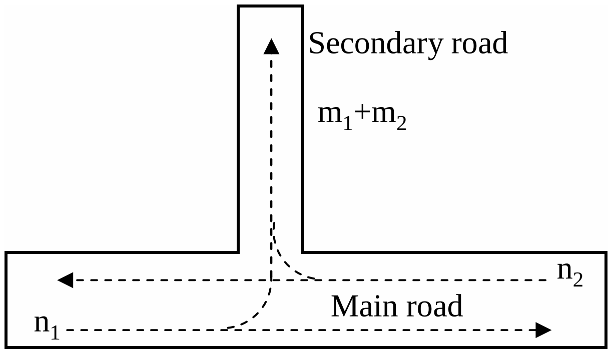
What they knew actually is the conditional probability of V given T : $P(V|T)$

$P(V) \neq P(V|T)$ meaning V and T are not independent!

Exercise 2.3

The observation of the traffic flow at a crossing, see the following figure, shows that $n_1 = 50$ vehicles move on the main road in direction 1. From those $m_1 = 25$ vehicles turn to the secondary road. $n_2 = 200$ vehicles move on the main road in direction 2 and $m_2 = 40$ vehicles turn to the secondary road.

How large is the probability that a vehicle moving on the main road will turn to the secondary road?



n_1 = 50 vehicles move on the main road in direction 1.
m_1 = 25 vehicles turn to the secondary road.
n_2 = 200 vehicles move on the main road in direction 2
m_2 = 40 vehicles turn to the secondary road.

Answer 2.3

Easy!! (Classical definition of probability)

$$P(B) = \frac{m_1 + m_2}{n_1 + n_2} = \frac{25 + 40}{50 + 200} = 0.26$$

Number of equally likely ways by which an experiment leads to B

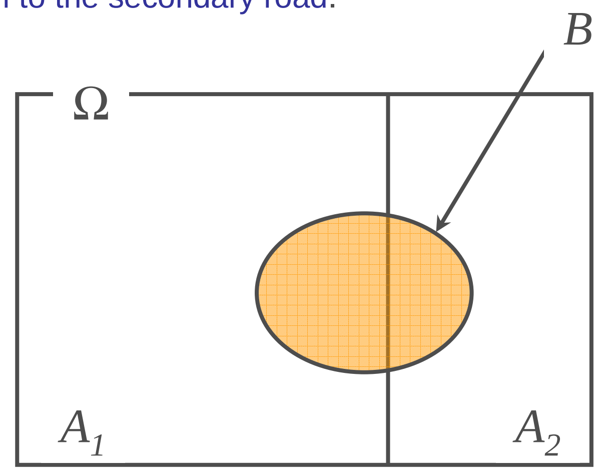
Total number of equally likely ways of the experiment

where *B* is the event that a vehicle turns to the secondary road.

BUT, think it a little more…..

$n_1 = 50$ vehicles move on the main road in direction 1.
$m_1 = 25$ vehicles turn to the secondary road.
$n_2 = 200$ vehicles move on the main road in direction 2
$m_2 = 40$ vehicles turn to the secondary road.

Answer 2.3

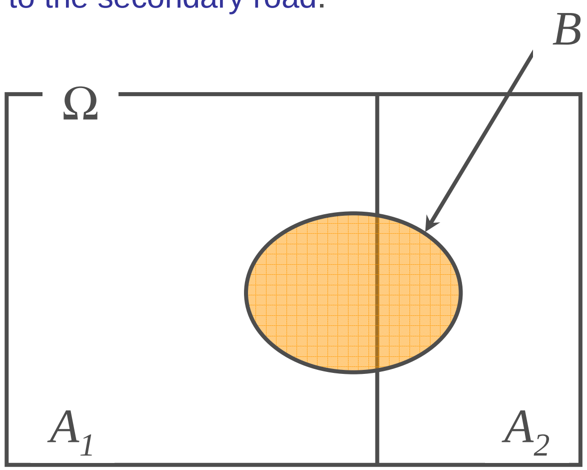


B is the event that a vehicle turns to the secondary road.

$$P(B) = P(B \cap A_1) + P(B \cap A_2) \Rightarrow$$

Answer 2.3

n_1 = 50 vehicles move on the main road in direction 1.
m_1 = 25 vehicles turn to the secondary road.
n_2 = 200 vehicles move on the main road in direction 2
m_2 = 40 vehicles turn to the secondary road.



B is the event that a vehicle turns to the secondary road.

$$P(B) = P(B \cap A_1) + P(B \cap A_2) \Rightarrow$$

$$P(B) = P(A_1) \cdot P(B | A_1) + P(A_2) \cdot P(B | A_2)$$

Probability that a vehicle is moving in one direction

$$P(A_1) = \frac{n_1}{(n_1+n_2)} = \frac{50}{(50+200)} = 0.2 \qquad P(A_2) = \frac{n_2}{(n_1+n_2)} = \frac{200}{(50+200)} = 0.8$$

Probability that a vehicle will turn to the secondary road

$$P(B | A_1) = \frac{m_1}{n_1} = \frac{25}{50} = 0.5 \qquad P(B | A_2) = \frac{m_2}{n_2} = \frac{40}{200} = 0.2$$

$$P(B) = P(A_1) \cdot P(B | A_1) + P(A_2) \cdot P(B | A_2) = 0.2 \cdot 0.5 + 0.8 \cdot 0.2 = 0.26$$

Exercise 2.4

Measurements are to be carried out with measurement devices.
Since a large number of devices is required, 20% of them will be provided by IAC (Institute for the Atmosphere and Climate) and 80% will be provided by IHW (Institute for Hydraulics and Water management).
5% of the devices provided by IAC do not fulfill the required accuracy, while 2% of the devices provided by IHW do not fulfill the required accuracy.

A student carried out a measurement using a device without knowing from which institute the device was provided. Thereby, she found the inaccuracy involved in the measurement.

How large is the probability that the measurement was carried out with a device provided by IAC?

20% of the devices will be provided by IAC
80% will be provided by IHW

5% of the devices provided by IAC do not fulfill the required accuracy
2% of the devices provided by IHW do not fulfill the required accuracy.

We do not know from which institute the device was provided.

How large is the probability that the measurement was carried out with a device provided by IAC?

Simplify the statement in the exercise!

A = Device of Institute A (IAC)
B = Device of Institute B (IHW)
D = Inaccurate device

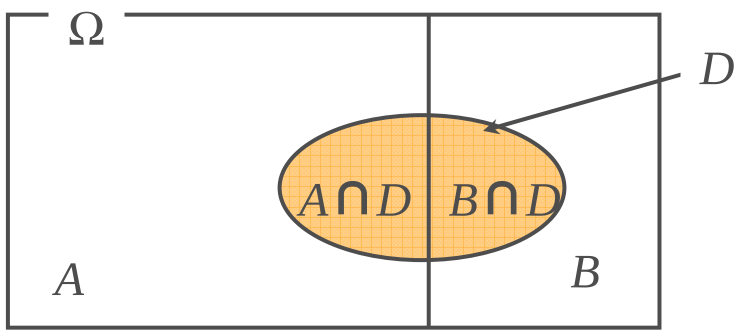
What is known???

$P(A) = 0.2 \qquad P(B) = 0.8$

Probability of using an inaccurate device:

$P(D \mid A) = 0.05 \qquad P(D \mid B) = 0.02$

What is required???



Exercise 2.5

A non destructive test method is carried out to determine whether the reinforcement of a component is corroded or not. From a number of past experiments, it is known that the probability that the reinforcement is corroded is 1%. If the reinforcement is corroded, this will be always indicated by the test. However, there is a 10% probability that the test will indicate that the reinforcement is corroded even if there is no corrosion.

How large is the probability that corrosion is present, if the non destructive test indicates corrosion? Calculate the probability using the Bayes' theorem.

Exercise 2.5 – Steps for solving

Identify the events:

K reinforcement is corroded

A test indication

What is known???

the probability that the reinforcement is corroded is 1%.

If the reinforcement is corroded, this will be always indicated by the test.

there is a 10% probability that the test will indicate that the reinforcement is corroded even if there is no corrosion

Exercise 2.5 – Steps for solving

Identify the events:

K reinforcement is corroded

A test indication

What is known???

the probability that the reinforcement is corroded is 1%.

If the reinforcement is corroded, this will be always indicated by the test.

there is a 10% probability that the test will indicate that the reinforcement is corroded even if there is no corrosion

What is required???

the probability that corrosion is present, if the non destructive test indicates corrosion

Exercise 2.6

The failure of a building in the city of Tokyo may be caused by two independent events:

F_1 : A big earthquake.

F_2 : A strong typhoon.

The annual probabilities of occurrence of the above events are:

$$P(F_1) = 0.04$$

$$P(F_2) = 0.08$$

Calculate the annual failure probability of the building.

Exercise 2.6

Identify the events:

F_1 : A big earthquake.

F_2 : A strong typhoon.

What is known?

$$P(F_1) = 0.04$$

$$P(F_2) = 0.08$$

What is required?

Calculate the annual failure probability of the building.

Exercise 2.7- (Group Exercise to be presented on 05.04.07)

Due to the increasing demand on drinking and processing water, the groundwater discharge flow has to be discussed. The hazard of long-term ground-lowering is analysed, whereby it is assumed that the ground-lowering depends on the thickness of the clay layer, h.
The thickness of the clay layer is classified in the following:

$C_1: \ 0 \leq h \leq 20cm$ $\qquad C_2: \ 20cm < h \leq 40cm$ $\qquad C_3: 40\,cm < h$

Based on experience a geologist estimates the following prior probabilities that the thickness of the clay layer at a site belongs to one of the above cases:

$P(C_1) = 0.2$ $\qquad P(C_2) = 0.47$

A geo-electrical test may be useful to update the prior probability on the ground category, although the test result may not always be correct. From past experience, the probabilities of the correct/false indication of the test are known as are listed in the (uncompleted) table below:

Category of thickness of clay layer C_i	Indication of the category of the thickness of the clay layer		
	$I = C_1$	$I = C_2$	$I = C_3$
C_1	0.84		0.03
C_2	0	0.77	
C_3		0.02	0.89

Exercise 2.7- (Group Exercise to be presented on 05.04.07)

Lots of information----Simplify!!!

•the ground-lowering depends on the thickness of the clay layer, h.

•Classification of thickness layer: (**Events)**

$C_1: \quad 0 \le h \le 20cm$ $\qquad C_2: \quad 20cm < h \le 40cm$ $\qquad C_3: \quad 40\,cm < h$

•prior probabilities: **(Known probabilities)**

$P(C_1) = 0.2$ $\qquad P(C_2) = 0.47$ $\qquad P(C_3) = 0.33$

•test to update the prior probability on the ground category, the test result may not always be correct.

•probabilities of the correct/false indication of the test:

Category of thickness of clay layer C_i	Indication of the category of the thickness of the clay layer		
	$I = C_1$	$I = C_2$	$I = C_3$
C_1	0.84		0.03
C_2	0	0.77	
C_3		0.02	0.89

Exercise 2.7- (Group Exercise to be presented on 05.04.07)

Lots of information----Simplify!!!

•the ground-lowering depends on the thickness of the clay layer, h.

•Classification of thickness layer: (**Events**) $C_1: \; 0 \le h \le 20cm$ $\quad C_2: \; 20cm < h \le 40cm$ $\quad C_3: \; 40\,cm < h$

•prior probabilities: (**Known probabilities**) $P(C_1) = 0.2$ $\quad P(C_2) = 0.47$ $\quad P(C_3) = 0.33$

•test to update the prior probability on the ground category, the test result may not always be correct.

•probabilities of the correct/false indication of the test:

Category of thickness of clay layer C_i	Indication of the category of the thickness of the clay layer		
	$I = C_1$	$I = C_2$	$I = C_3$
C_1	0.84		0.03
C_2	0	0.77	
C_3		0.02	0.89

What is required???

a. Complete the table

b. A geo-electrical test was carried out and indicated C_3 as the thickness of the clay layer. What is the probability that the thickness of the clay layer belongs to C_1 , C_2 , C_3 ?

Exercise 2.7- (Group Exercise to be presented on 05.04.07)

What is required???

a. Complete the table
b. A geo-electrical test was carried out and indicated C_3 as the thickness of the clay layer. What is the probability that the thickness of the clay layer belongs to C_1, C_2, C_3 ?

a. Complete the table

Category of thickness of clay layer C_i	Indication of the category of the thickness of the clay layer		
	$I = C_1$	$I = C_2$	$I = C_3$
C_1	0.84		0.03
C_2	0	0.77	
C_3		0.02	0.89

$$P(I = C_1 | C_1) + P(I = C_2 | C_1) + P(I = C_3 | C_1) = 1$$

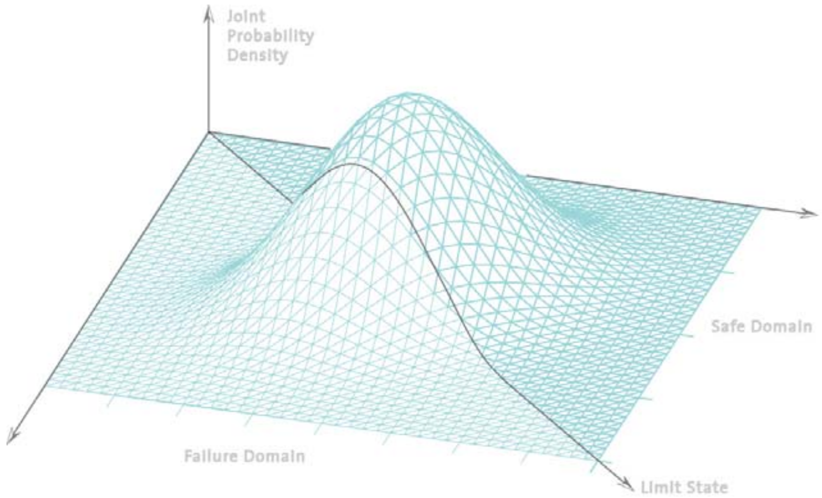
Exercise 2.7- (Group Exercise to be presented on 05.04.07)

What is required???

a. Complete the table
b. A geo-electrical test was carried out and indicated C_3 as the thickness of the clay layer. What is the probability that the thickness of the clay layer belongs to C_1 , C_2 , C_3 ?

b. **How can we express this??**

Use Bayes' theorem (script section B.5)


Joint
Probability
Density
Safe Domain
Failure Domain
Limit State

QUESTIONS ? ? ? ? ? ?