Statistics primer for Bayesian classifiers Lecture 4.

BOOLEAN VALUED RANDOM VARIABLES

Discrete Boolean-valued random variables

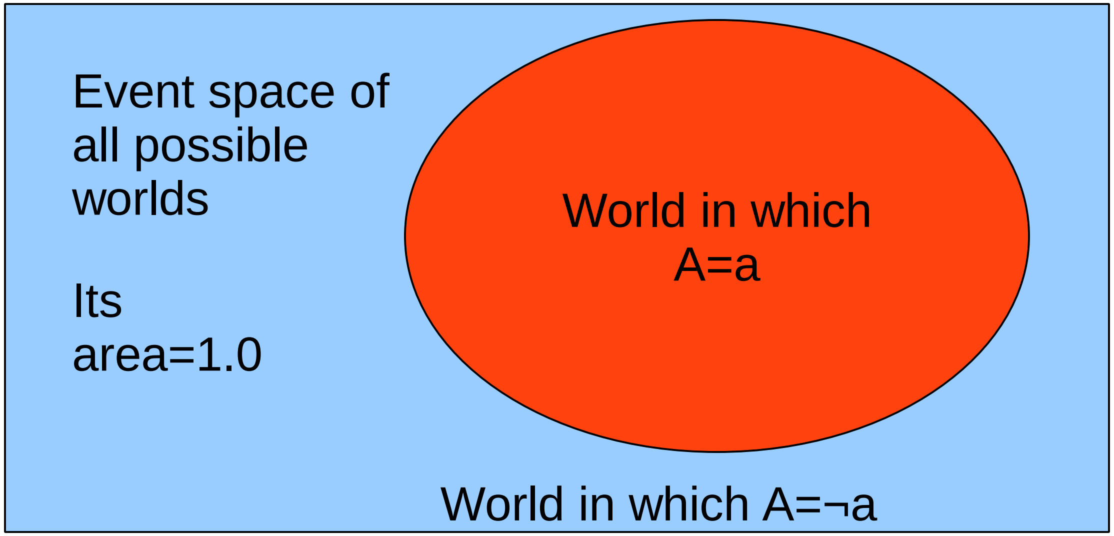
A is a Boolean-valued random variable if *A* denotes an event, and there is some degree of uncertainty as to whether *A* occurs or not.

Examples:

- P = p: The US president in 2023 will be male
- P=¬p: The US president will not be a male
- H = h: You wake up tomorrow with a headache
- H=¬h: No headache

Probabilities

We write P(A=a), or P(A=true) or simple P(A) as "the fraction of possible worlds where A=a is true"

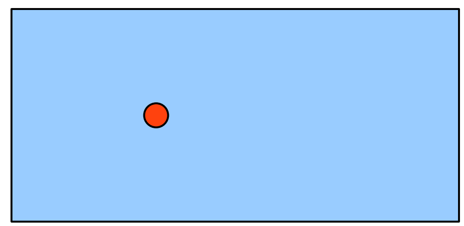


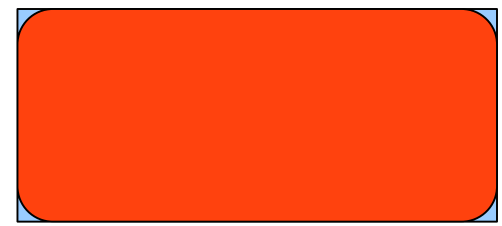
P(A=a) is the proportion of a red oval out of the blue universe

The Axioms of Probability

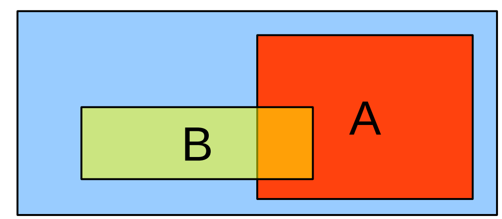
We do not need to prove that:

I. 0<= P(A=a)<=1

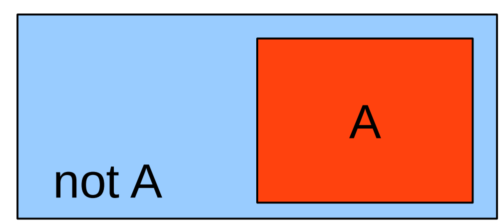




II. P(A or B)=P(A)+P(B)-P(A and B)

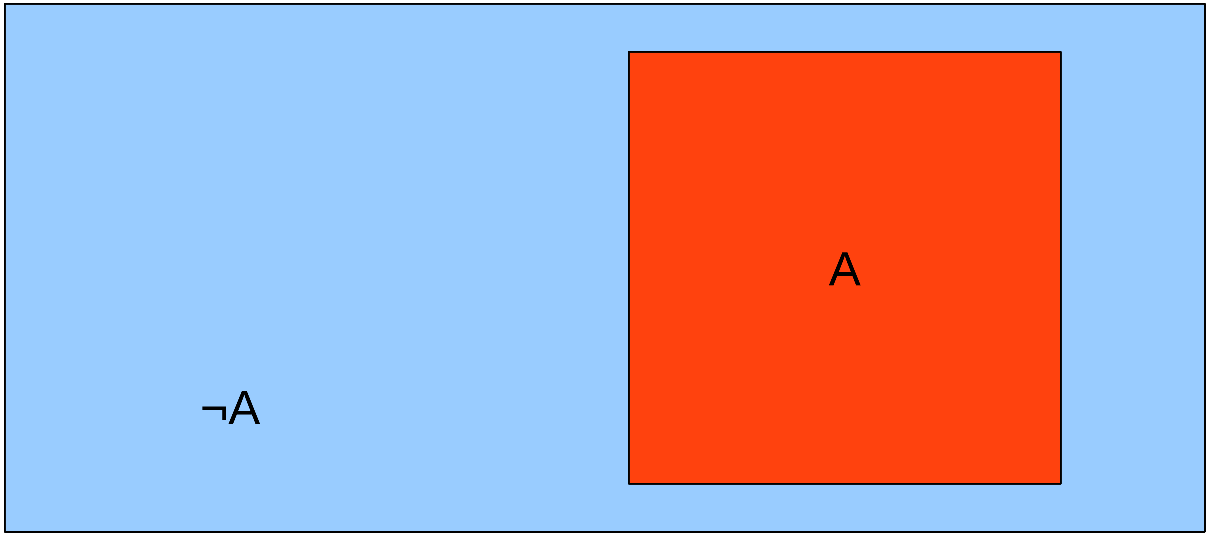


III. P(A)+P(¬A)=1



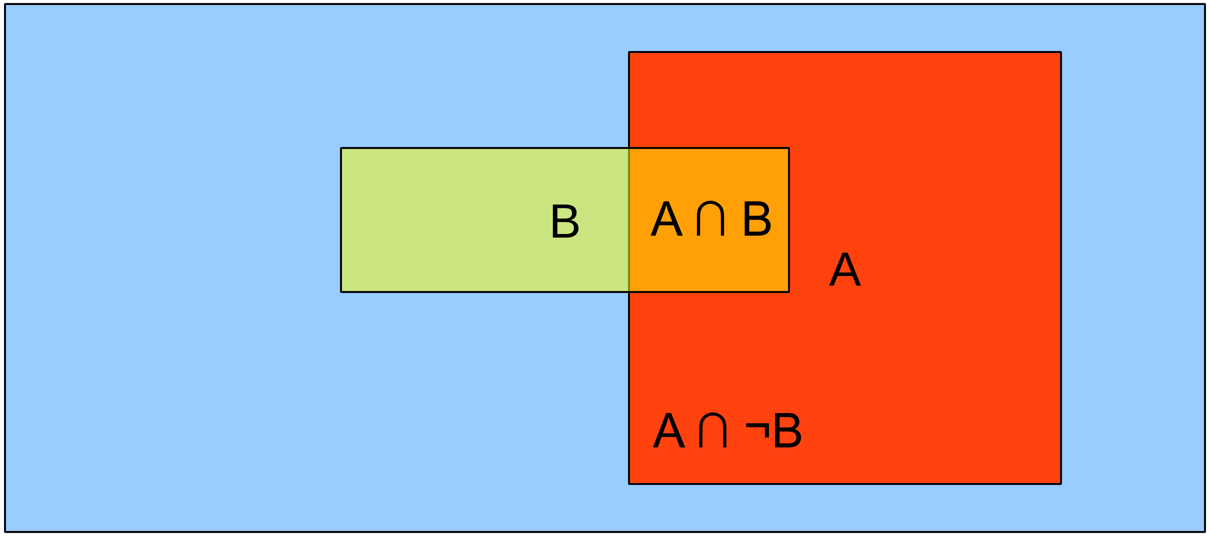
Theorems of Probability I

$P(\neg A) = 1 - P(A)$



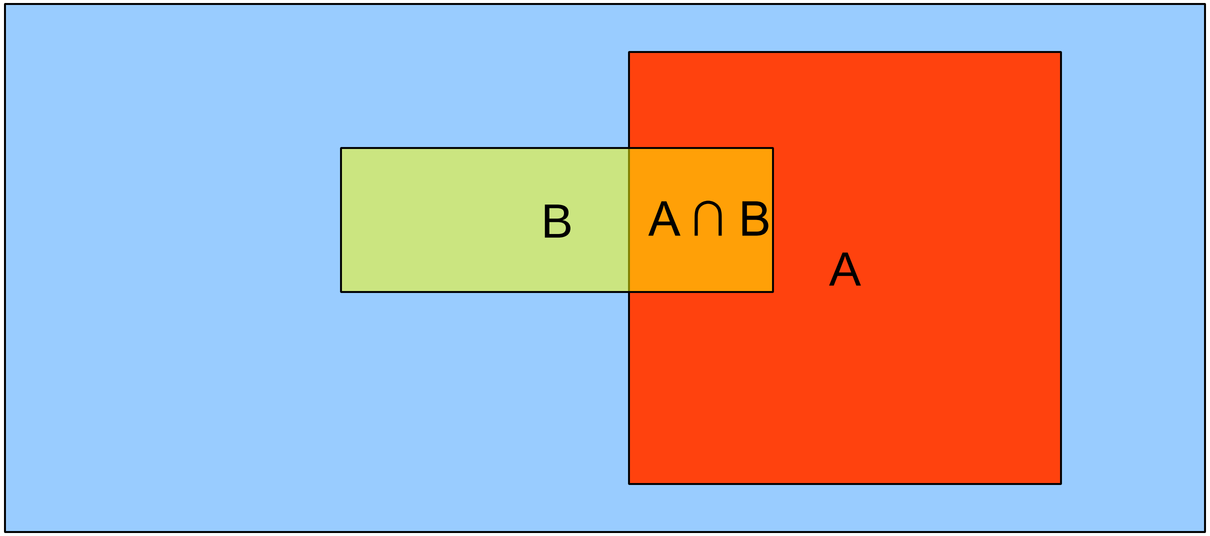
Theorems of Probability II

$P(A)=P(A \cap B) + P(A \cap \neg B)$



Conditional probability I

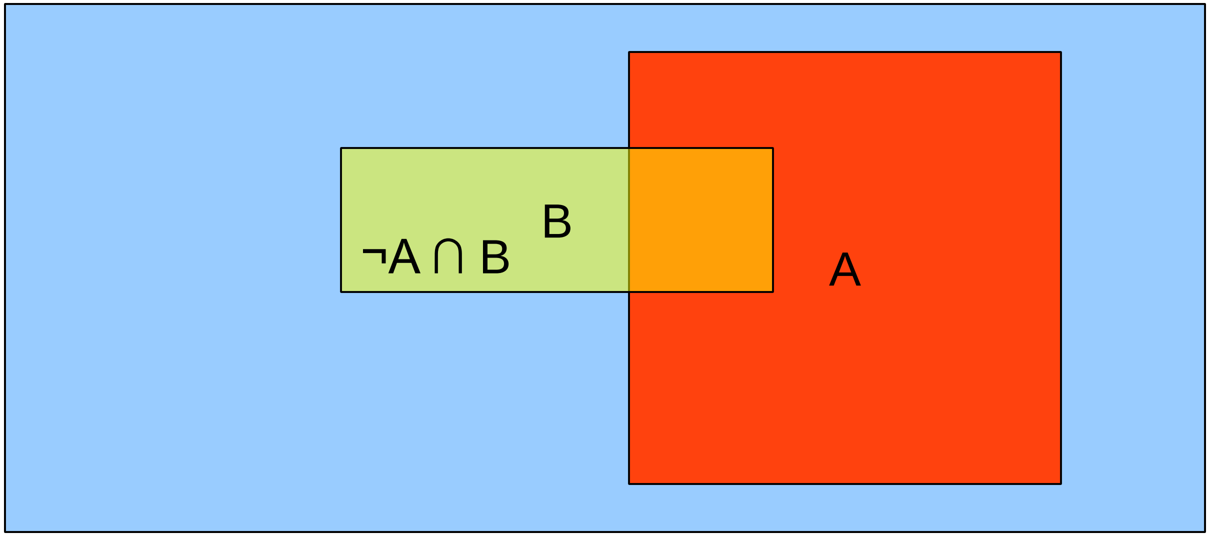
- P(A|B) = fraction of worlds in which A is true out of all the worlds where B is true



CP definition: $P(A|B) = P(A \cap B) / P(B)$

Conditional probability II

- P(A|B) = fraction of worlds in which A is true out of all the worlds where B is true



CP definition: $P(\neg A|B) = P(\neg A \cap B) / P(B)$

Probabilistic independence

Two random variables A and B are independent if

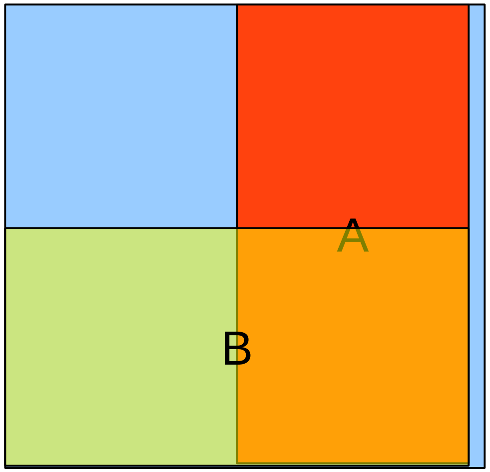
P(A|B) = P(A), which means that:

P(a|b) = P(a)

P(¬ a|b) = P(¬ a)

P(a| ¬ b) = P(a)

P(a| ¬ b) = P(a)



Knowing that B is true (or false) does not change the probability of A

Theorems III. Chain rule

From the definition of conditional probabilities:

$$P(A|B) = P(A \cap B) / P(B)$$

we can compute $P(A \cap B)$ – that both events happened together:

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

If A and B are independent:

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

Theorems IV. Bayes theorem

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

On the other hand:

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

$$P(A|B)P(B) = P(B|A)P(A)$$

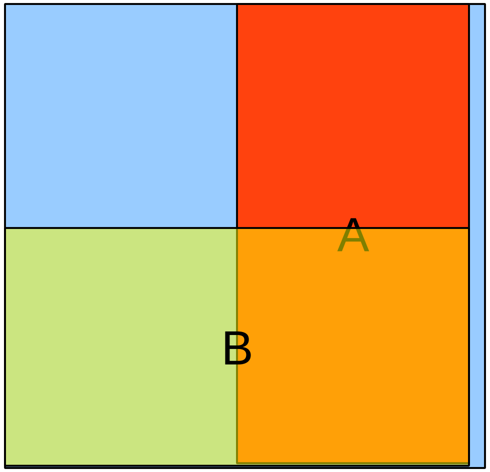
and we can express conditional probability of A given B through conditional probability of B given A and unconditional probabilities of A and B:

$$P(A|B) = P(B|A)P(A)/P(B)$$

Independent and mutually exclusive events

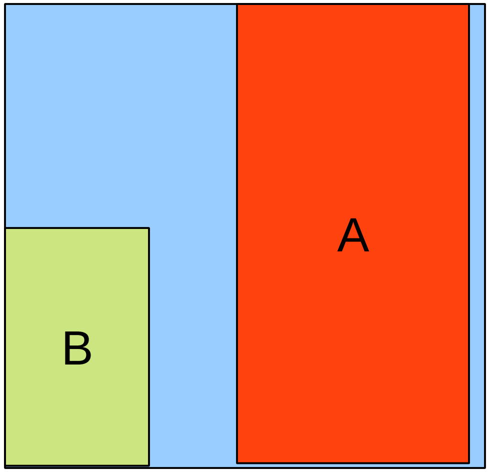
A is *independent* of B: knowing that B is true (or false) does not change the probability of A:

$$P(A|B) = P(A)$$



A and B are *mutually exclusive* – not independent variables: if A is true then B is false, if A is false then B is true with probability $P(B|\neg A)$

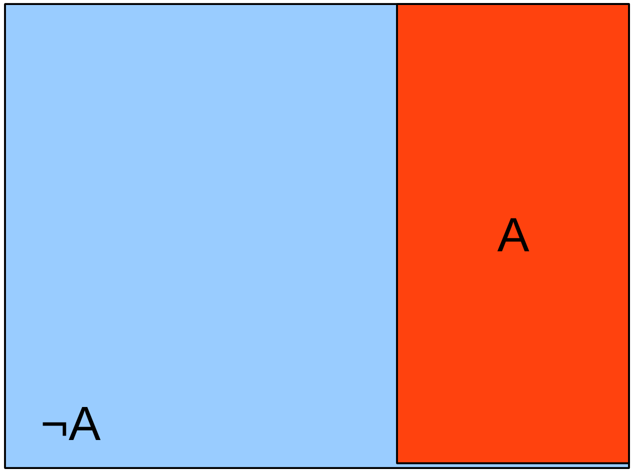
$$P(A \cap B)=0$$



Theorems of Probability V

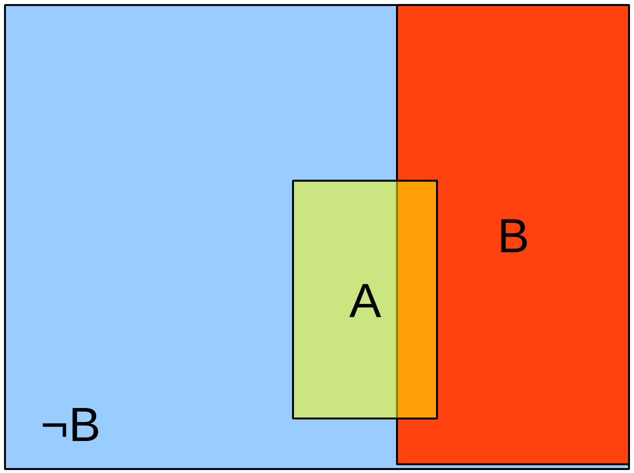
A and ¬A are mutually exclusive, so Axiom II becomes:

$P(A \vee \neg A) = P(A) + P(\neg A)$



Theorems of Probability VI

P(A ∩ (B ∨ ¬B))=P(A ∩ B) +P(A ∩ ¬B)=P(A)
(from Theorem II)



Multiple variables

The theorems for 2 Boolean-valued random variables can be extended to several random variables *C, E1, E2,…,En.* Let C, E1, E2, … En be Boolean-valued random variables. For convenience, we will let E denote the n-tuple of random variables (E1,E2,…,En)

E1, E2, … En=E

P(C ∩ E1 ∩ E2 ∩ … ∩ En)=P(C,E1,E2,…En)=P(C,E)

Chain rule:

$P(C,E)=P(C)P(E_1|C,E_2,\ldots E_n)P(E_2|C,E_1, E_3,\ldots,E_n)x\ldots xP(E_n|C,E_1,\ldots E_{n-1})$

Multiple variables

If E1,…En are mutually independent and depend only on C then:

$P(C,E)=P(C)P(E_1|C)P(E_2|C)x\ldots xP(E_n|C)$

Bayes theorem:

$P(C|E)=P(C,E)/P(E)$

Multi-valued random variables

Suppose A can take a value from a set of size greater than 2 – say, k value. *Multi-valued* random variable is defined as:

- $P(A=ai \cap A=aj)=0$ for $i \neq j$ (mutually exclusive)
- $P(A=a1 \vee A=a2 \vee \ldots \vee A=ak)=1$

Theorem V: $P(A=a1 \vee A=a2 \vee \ldots A=am)=\Sigma_{(\text{from } i=1 \text{ to } m)}P(A=ai)$, $m<=k$

Theorem VI: $P(B \cap [A=a1 \vee A=a2 \vee A=am])=\Sigma_{(\text{from } i=1 \text{ to } m)} P(B \cap Ai)$