Statistics primer
for Bayesian thinking

BOOLEAN VALUED RANDOM VARIABLES

Discrete Boolean-valued random variables

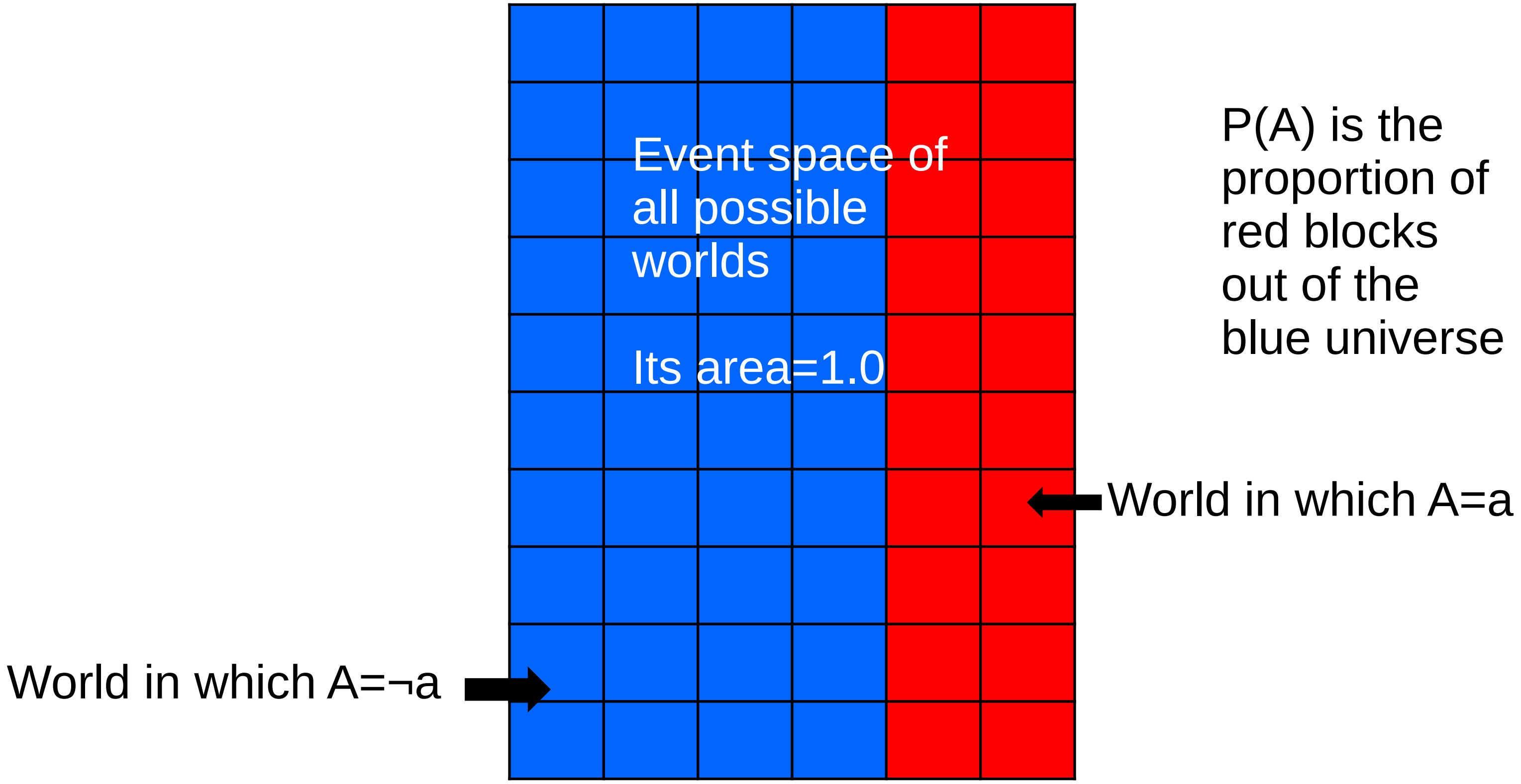
A is a Boolean-valued random variable if *A* denotes an event, and there is some degree of uncertainty as to whether *A* occurs or not.

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”

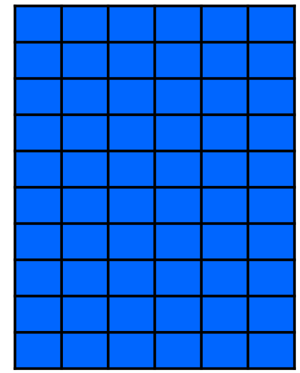


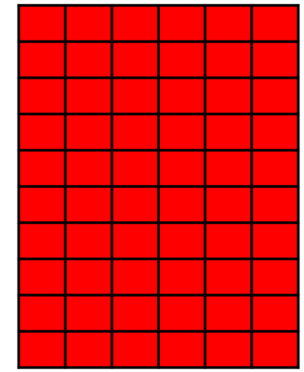
P(A) is the proportion of red blocks out of the blue universe

The Axioms of Probability

We do not need to prove that:

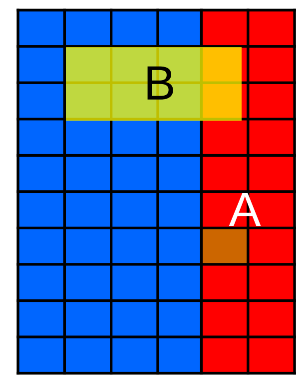
I. 0<= P(A=a)<=1



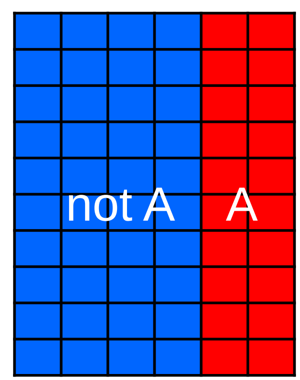


II. P(A or B)=P(A)+P(B)-P(A and B)

A B

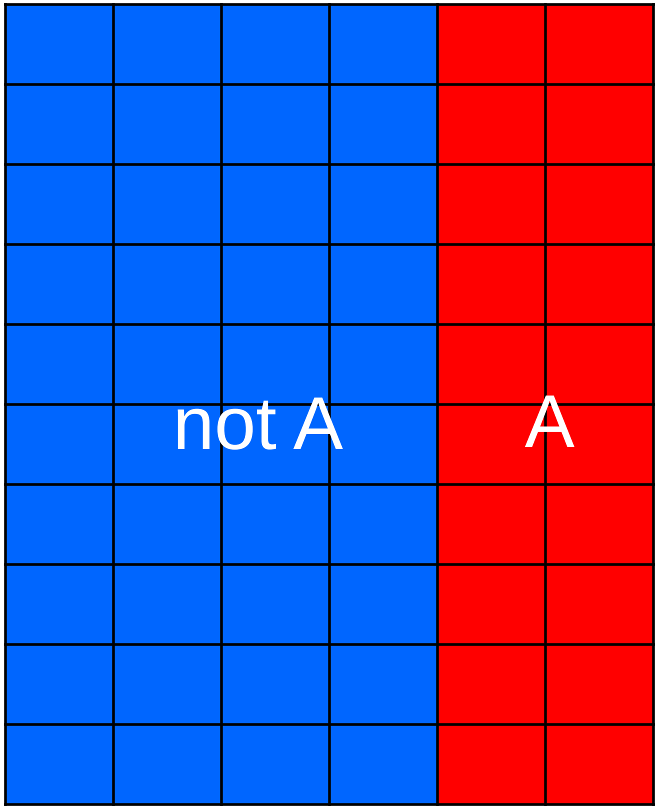


III. P(A)+P(¬A)=1



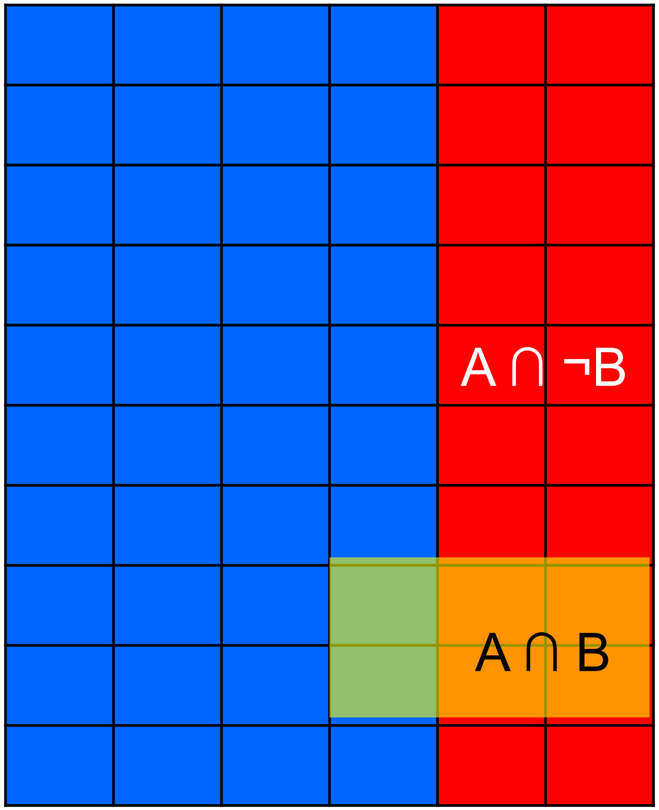
Theorems of Probability: Theorem 1

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



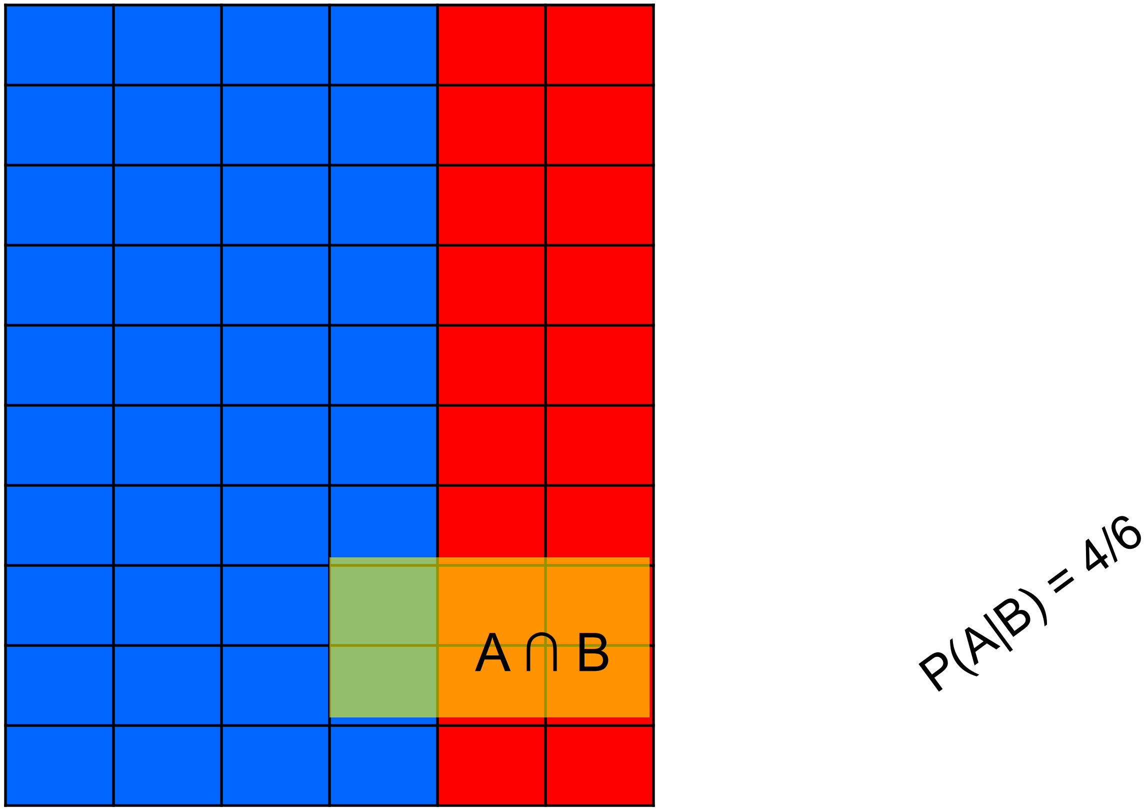
Theorems of Probability: Theorem 2

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



Conditional probability: definition

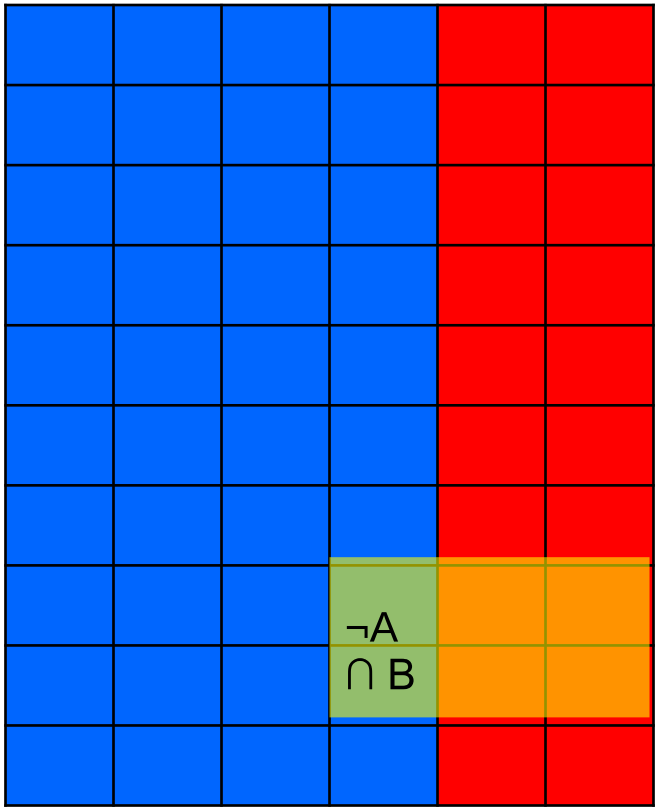
- P(A|B) = fraction of worlds in which A is true out of all the worlds where B is true



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

Conditional probability: definition

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

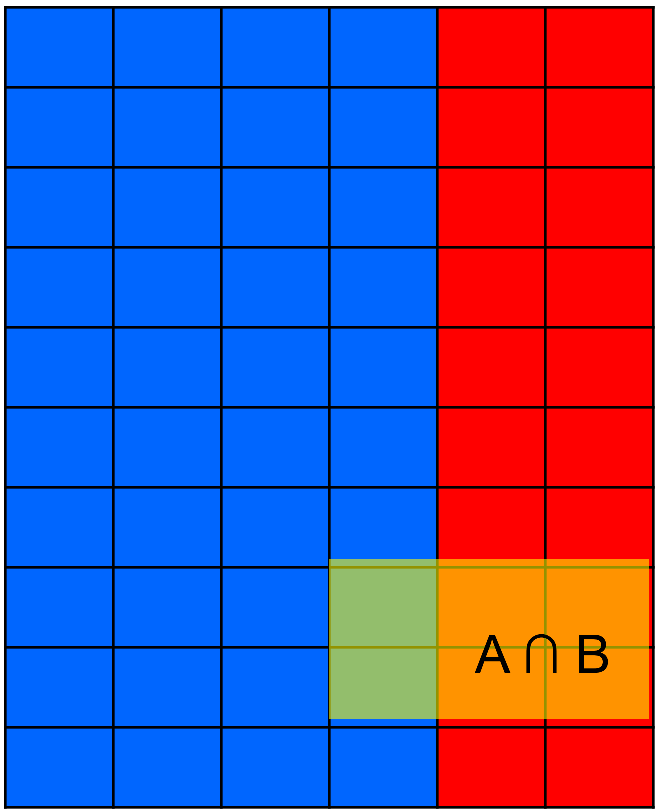


$P(\neg A|B) = 2/6$

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

Conditional probability: definition

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



P(B|A) = 4/20 =0.2
P(B) = 6/60 = 0.1

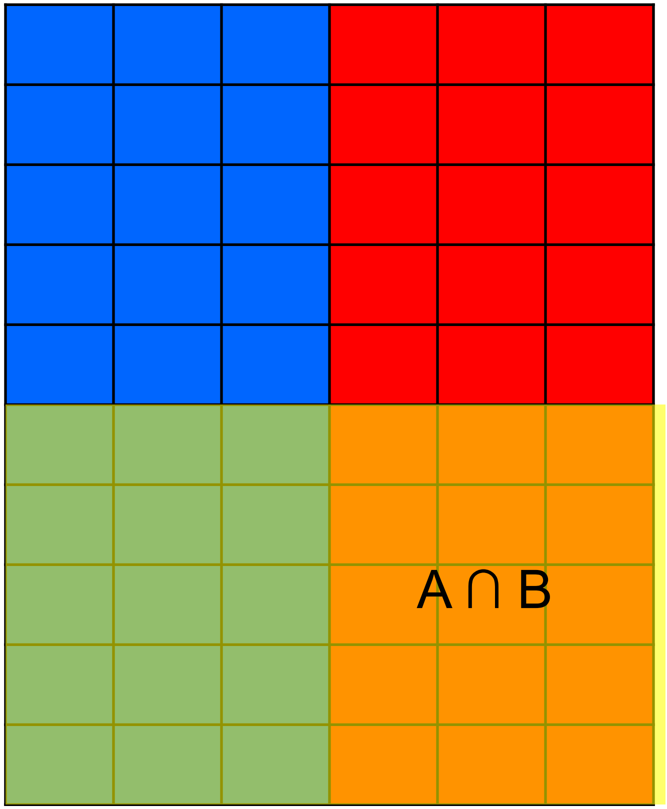
$P(B|A) = P(A \cap B) / P(A)$

Probabilistic independence

Two random variables A and B are *mutually independent* if

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

P(A|B) = P(A)	15/30 = 30/60
P(¬ A|B) = P(¬ A)	15/30 = 30/60
P(A| ¬ B) = P(A)	15/30 = 30/60
P(¬A| ¬ B) = P(A)	15/30 = 30/60



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

Theorems 3. 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* that becomes:

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

Theorems 4. Bayes theorem

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

From definition of Conditional probability

On the other hand:

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

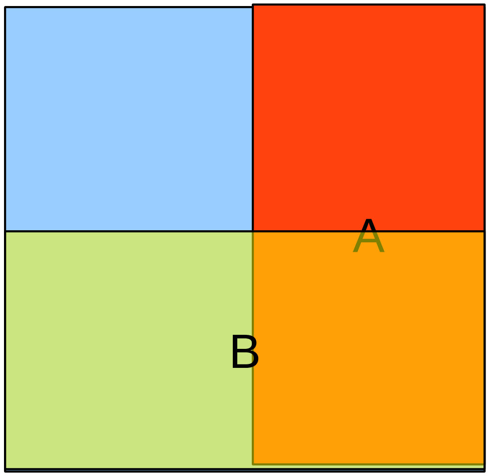
$$\Downarrow$$

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

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

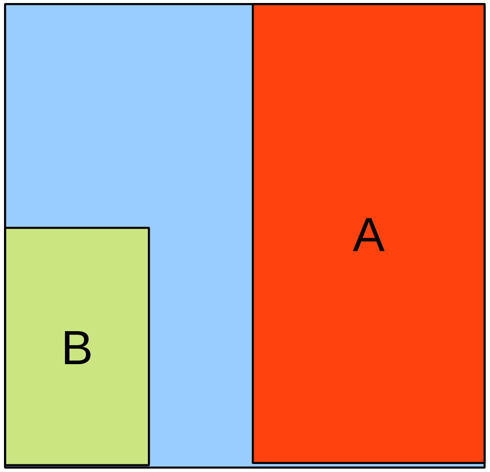
$$P(A|B) = P(B|A)P(A)/P(B)$$

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 $P(B|\neg A)$

$P(A \cap B)=0$

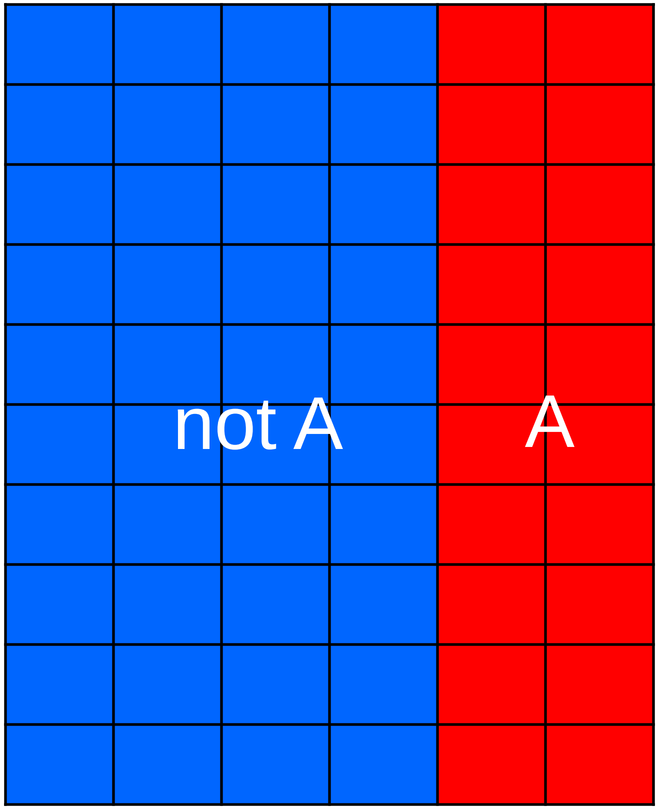
Theorems of Probability 5

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

P(A or B)=P(A)+P(B)-~~P(A and B)~~

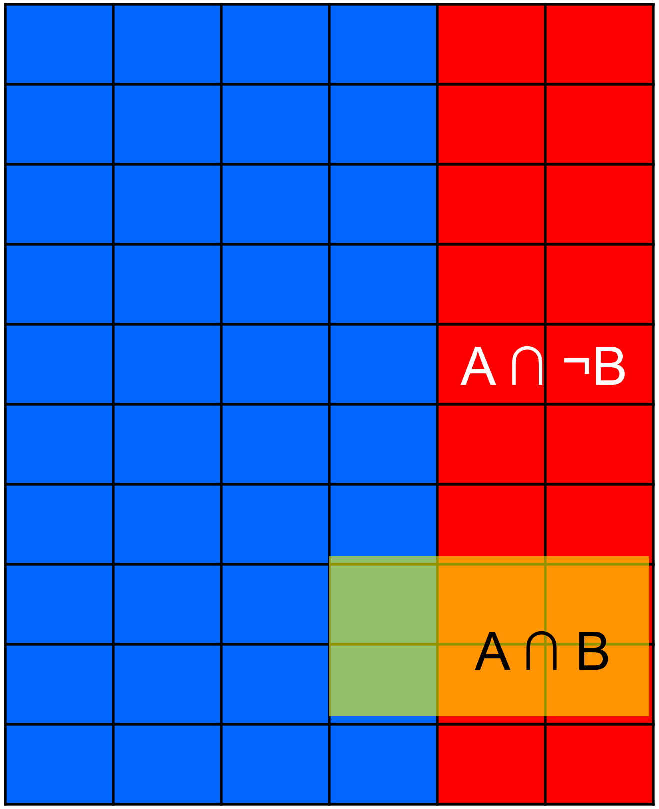
becomes:

P(A or ¬A)=P(A) + P(¬A) = 1



Theorems of Probability 6

P(A ∩ (B or ¬B))=P(A)



Using Bayes theorem for diagnosis

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

Let *C* be a random variable which represents the probability of some condition to be True (you have Pneumonia)

We can consider a general probability of the population to have pneumonia – known from statistical data: P(*C*) – *prior probability*

But you have a symptom – cough. We want to know P(C|E), where E is evidence – symptom

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

Most of the times P(E/C) is known, and P(E) is either known or we can get away without it

Multiple Boolean random variables

All theorems for 2 Boolean-valued random variables can be extended to several random variables $C, E_1, E_2,\ldots,E_n$.

Let $C, E_1, E_2,\ldots,E_n$ be Boolean-valued random variables.

For convenience, we will let E denote the n-tuple of random variables $(E_1, E_2,\ldots,E_n)$

$$E_1, E_2,\ldots,E_n = E$$

$$P(C \cap E_1 \cap E_2 \cap \ldots \cap E_n)=P(C,E_1,E_2,\ldots E_n)=P(C,E)$$

Just a notation

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 dependent on C

C – condition
E – evidence (event)

If $E_1,\ldots E_n$ are mutually independent and depend only on *C* then:

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

And from Bayes theorem:

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

That gives you a formula of the probability that the unknown condition C was true given a set of known evidences E

Multi-valued random variables

Suppose A is not a Boolean variable but can take a value from a set of size greater than 2 – say, *k* values. *Multi-valued* random variable is defined as:

- $P(A=a_i \cap A=a_j)=0$ for $i \neq j$ (mutually exclusive)
- $P(A=a_1 \text{ or } A=a_2 \text{ or } \ldots \text{ or } A=a_k)=1$

Theorem 5 becomes:

$P(A=a_1 \text{ or } A=a_2 \text{ or } \ldots A=a_m)=\Sigma_{(\text{from } i=1 \text{ to } m)}P(A=a_i)$, $m<=k$

Theorem 6 becomes:

$P(B \cap [A=a_1 \text{ or } A=a_2 \text{ or } \ldots A=a_m])=\Sigma_{(\text{from } i=1 \text{ to } m)} P(B \cap A_i)$