Probability

Probability in Computing Science

Almost any advance computing application today has some randomization or statistical :

- network security
- cryptography
- web search and web advertising (ranking results or adverts)
- spam filtering
- social network tools
- recommendation systems: Netflix, Amazon, Google, ...
- communication protocols
- computational finance
- system biology
- DNA sequencing and analysis
- data mining

Probability in Computing Science

Randomized algorithms

- cryptography and security
- fast algorithms
- simulations

Probabilistic analysis of algorithms

Statistical inference

- machine learning, deep learning, data mining...

All are based on the same (mostly discrete) probability theory principles and techniques

Before we go on...

- The definitions, axioms and formulas we will encounter next are **very** tightly relalated to the concept of
  - Sets
  - Cardinality of sets (number of elements in a set)
    - Especially cardinality of union of sets
  - Venn diagrams

- Stop and revisit these topics if you think you need to.

A Quick Overview of Probability - Outline

- Probability – Introduction & Definitions
- Axioms and Properties of Probability
- Conditional Probability

First: What is a “Proposition”

- It is a “declarative” sentence that is either TRUE or FALSE (but not both).

- More about this later…

Probability – Introduction

In the real world, we often do not know whether a given *proposition* is `true` or `false`

Probability theory gives us a way to reason about propositions whose truth is **uncertain.**

Probability – Introduction

Probability theory deals with such (random) experiments whose truth value is uncertain

- these are processes or actions whose outcome cannot be predicted with certainty and may differ if the experiment is repeated

Probability – Introduction

Probability theory deals with (random) experiments

- these are processes or actions whose outcome cannot be predicted with certainty and may differ if the experiment is repeated

Examples:

- rolling a dice
  - possible outcomes: $1, \ldots, 6$

Probability – Introduction

Probability theory deals with (random) experiments

- these are processes or actions whose outcome cannot be predicted with certainty and may differ if the experiment is repeated

Examples:

- rolling a dice
  - possible outcomes: 1,…,6

- drawing two cards from a shuffled pack of cards
  - possible outcomes are pairs of cards: {♣5,♦6},{♠J,♣A},…

Probability – Introduction

Probability theory deals with (random) experiments

- these are processes or actions whose outcome cannot be predicted with certainty and may differ if the experiment is repeated

Examples:

- rolling a dice
  - possible outcomes: 1,…,6

- drawing two cards from a shuffled pack of cards
  - possible outcomes are pairs of cards: {♣5,♦6},{♠J,♣A},…

- flipping a coin
  - possible outcomes: heads and tails

Probability – Introduction

Probability theory deals with (random) experiments

- these are processes or actions whose outcome cannot be predicted with certainty and may differ if the experiment is repeated

Examples:

- rolling a dice
  - possible outcomes: 1,...,6

- drawing two cards from a shuffled pack of cards
  - possible outcomes are pairs of cards: {♣5,♦6},{♠J,♣A},...

- flipping a coin
  - possible outcomes: heads and tails

Probability theory aims at quantifying the uncertainty surrounding the possible outcomes of an experiment

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

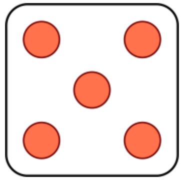
E.g. Rolling a dice

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the *world*



E.g. result of rolling a dice

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

E.g for rolling a dice

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

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

E.g. $A = \{2, 4, 6\}$

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Frequency: how often an outcome occurs in a sequence of experiments

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset A⊆Ω of possible outcomes with some common property

Frequency: how often an outcome occurs in a sequence of experiments

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

Eg, Given: $A = \{2, 4, 6\}$

Then $P(A) = 0.5$

Calculating Probability

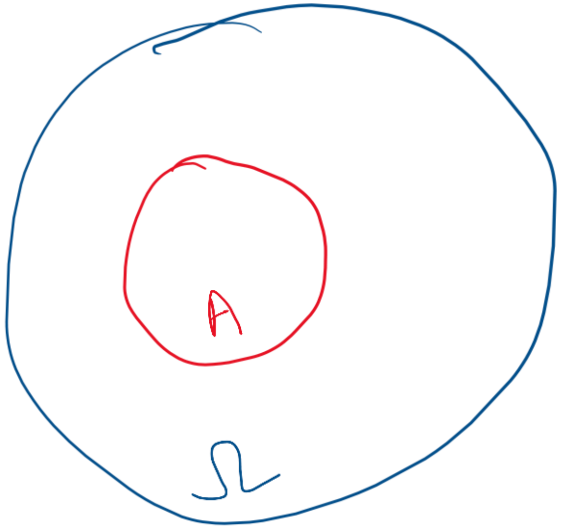
If Ω is a finite **sample space…**

of equally likely **outcomes**,

and A is an **event**, *(that is, a subset of* Ω), then

the ***probability*** of *A* is

$$P(A) = \frac{|A|}{|\Omega|}$$



Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

- P = 0 → This event will never happen
- P = 1 → This event will always happen
- P = 0.5 → There is a 50% chance of this event happening

Calculating Probability

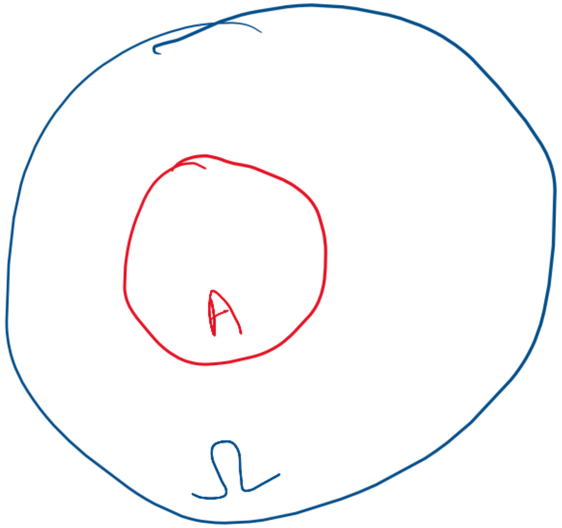
If Ω is a finite **sample space…**

of equally likely **outcomes**,

and A is an **event**, *(that is, a subset of* Ω), then

the ***probability*** of A is

$$P(A) = \frac{|A|}{|\Omega|}$$



Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment
- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen
P = 1 → This event will always happen
P = 0.5 → There is a 50% chance of this event happening

Probability – Definitions

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

– one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset A⊆Ω of possible outcomes with some common property

Frequency: how often an outcome occurs in a sequence of experiments

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

Eg, Given: $A = \{2, 4, 6\}$

Then $P(A) = 0.5$

$$\boxed{P(A) = \frac{|A|}{|\Omega|}}$$

Probability – Examples

Rolling a single die 

- sample space $\Omega=\{\ 1,\ 2,\ 3,\ 4,\ 5,\ 6\ \}$
  - 6 possible outcomes all equally likely
- probability of rolling an even number: $\boxed{P(A) = \frac{|A|}{|\Omega|}}$

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment
- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

Probability – Examples

Rolling a single die 

- sample space Ω={ 1, 2, 3, 4, 5, 6 }
  - 6 possible outcomes all equally likely
- probability of rolling an even number:
  - A = {2, 4, 6}
  - P[A] = 3/6 = 1/2

$$P(A) = \frac{|A|}{|\Omega|}$$

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment
- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen
P = 1 → This event will always happen
P = 0.5 → There is a 50% chance of this event happening

Probability – Examples

Rolling a single die 

- sample space Ω={ 1, 2, 3, 4, 5, 6 }
  - 6 possible outcomes all equally likely
- probability of rolling an even number:
  - **A** = {2, 4, 6}
  - **P**[A] = 3/6 = 1/2
- probability of not rolling a 6

$$\boxed{P(A) = \frac{|A|}{|\Omega|}}$$

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use **P**[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

Probability – Examples

Rolling a single die 

- sample space Ω={ 1, 2, 3, 4, 5, 6 }
  - 6 possible outcomes all equally likely
- probability of rolling an even number:
  - **A** = {2, 4, 6}
  - **P**[A] = 3/6 = 1/2
- probability of not rolling a 6
  - A = {1,2,3,4,5}
  - P[A] = 5/6

$$\boxed{P(A) = \frac{|A|}{|\Omega|}}$$

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use **P**[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

$$P(A) = \frac{|A|}{|\Omega|}$$

Rolling two dice

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability that sum is even?

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

– one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset A⊆Ω of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

- P = 0 → This event will never happen
- P = 1 → This event will always happen
- P = 0.5 → There is a 50% chance of this event happening

$$\boxed{P(A) = \frac{|A|}{|\Omega|}}$$

Rolling two dice

– sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$

- $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely

– probability that sum is even → **using the product rule and sum rules**

– Why product and sum rules here?

- When computing probability we are essentially "counting" number of ways *something happens* (an event), that may be composed of:
  - multiple separate tasks that together form that event, e.g. two dice being thrown → product rule
  - different ways the same event can happen (an event can happen like this OR that → sum rule
- Thus product rule and sum rule find application here as well

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

- P = 0 → This event will never happen
- P = 1 → This event will always happen
- P = 0.5 → There is a 50% chance of this event happening

$$\boxed{P(A) = \frac{|A|}{|\Omega|}}$$

Rolling two dice

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability that sum is even (*using the product and sum rules*):
  - What is "A" here? It is all possible ways both dice are even (sum is even), OR both dice are odd (sum is even). So,

- Thus, $P(A) = \dfrac{\overbrace{ways\ both\ dice\ can\ be\ even + ways\ both\ dice\ can\ be\ odd}^{(sum\ rule)}}{Total\ number\ of\ ways\ two\ dice\ can\ be\ thrown}$

$$= \frac{\overbrace{(ways\ \mathbf{first}\ dice\ can\ be\ even \cdot ways\ \mathbf{second}\ dice\ can\ be\ even)}^{(product\ rule)} \overbrace{+}^{(sum\ rule)} \overbrace{(ways\ first\ dice\ can\ be\ odd \cdot ways\ second\ dice\ can\ be\ odd)}^{(product\ rule)}}{Total\ number\ of\ ways\ two\ dice\ can\ be\ thrown}$$

$$= (3 \cdot 3) + (3 \cdot 3)/36 = 1/2$$

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

$$P(A) = \frac{|A|}{|\Omega|}$$

Rolling two dice

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \land 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability that sum is even (*using the product and sum rules*):
  - What is "A" here? It is all possible ways both dice are even (sum is even), OR both dice are odd (sum is even). So,

(sum rule)

- Thus, P(A) = $\dfrac{\text{ways both dice can be even} + \text{ways both dice can be odd}}{\text{Total number of ways two dice can be thrown}}$

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset A⊆Ω of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

- P = 0 → This event will never happen
- P = 1 → This event will always happen
- P = 0.5 → There is a 50% chance of this event happening

$$\boxed{P(A) = \frac{|A|}{|\Omega|}}$$

Rolling two dice 

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability that sum is even (*using the product and sum rules*):
  - What is "A" here? It is all possible ways both dice are even (sum is even), OR both dice are odd (sum is even). So,

(sum rule)

- Thus, P(A) = $\dfrac{ways\ both\ dice\ can\ be\ even + ways\ both\ dice\ can\ be\ odd}{Total\ number\ of\ ways\ two\ dice\ can\ be\ thrown}$

(product rule) *(sum rule)* *(product rule)*

$$= \frac{(ways\ \mathbf{first}\ dice\ can\ be\ even \cdot ways\ \mathbf{second}\ dice\ can\ be\ even) + (ways\ first\ dice\ can\ be\ odd \cdot ways\ second\ dice\ can\ be\ odd)}{Total\ number\ of\ ways\ two\ dice\ can\ be\ thrown}$$

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

– one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset A⊆Ω of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

$$P(A) = \frac{|A|}{|\Omega|}$$

Rolling two dice

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability that sum is even (*using the product and sum rules*):
  - What is "A" here? It is all possible ways both dice are even (sum is even), OR both dice are odd (sum is even). So,
- Thus, P(A) =

$$\frac{\overbrace{\textit{ways both dice can be even} + \textit{ways both dice can be odd}}^{\text{(sum rule)}}}{\textit{Total number of ways two dice can be thrown}}$$

$$= \frac{\overbrace{(\textit{ways first dice can be even} \cdot \textit{ways second dice can be even})}^{\text{(product rule)}} \overbrace{+}^{\text{(sum rule)}} \overbrace{(\textit{ways first dice can be odd} \cdot \textit{ways second dice can be odd})}^{\text{(product rule)}}}{\textit{Total number of ways two dice can be thrown}}$$

$$= (3 \cdot 3) + (3 \cdot 3)/36 = 1/2$$

Probability – Examples

Experiment (or trial): an occurrence with an uncertain outcome

Outcome: the result of an experiment

- one particular state of the world

Sample space: the set Ω of all possible outcomes for the experiment

Event: subset $A \subseteq \Omega$ of possible outcomes with some common property

Probability: the probability of an event is the degree of certainty than an event will occur (use P[A] to denote probability of event A)

P = 0 → This event will never happen

P = 1 → This event will always happen

P = 0.5 → There is a 50% chance of this event happening

$$P(A) = \frac{|A|}{|\Omega|}$$

Rolling two dice

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability sum equals 7:

P[(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)] = 6/36

A Quick Overview of Probability – Outline

- Probability – Introduction & Definitions
- Axioms and Properties of Probability
- Conditional Probability

Probability – Axioms

There are three basic axioms of probability from which everything else can be derived. Given that:

- we have a sample space Ω and events $A, A \subseteq \Omega$
- the probability of the event A is denoted $\mathbf{P}[A]$
- events A and B are mutually exclusive if $A \cap B = \emptyset$

Axiom: a statement accepted as true as the basis for argument or inference

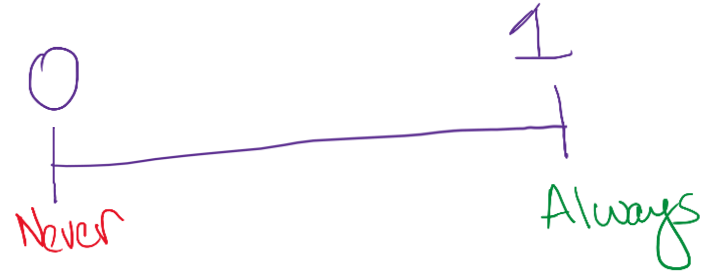
https://www.merriam-webster.com/dictionary/axiom

Probability – Axioms

There are three basic axioms of probability from which everything else can be derived. Given that:

- we have a sample space Ω and events $A, B \subseteq \Omega$
- the probability of the event A is denoted $\mathbf{P}[A]$
- events A and B are mutually exclusive if $A \cap B = \emptyset$

Axiom1: $\mathbf{0 \leq P[A] \leq 1}$ **for all events** $\mathbf{A \subseteq \Omega}$

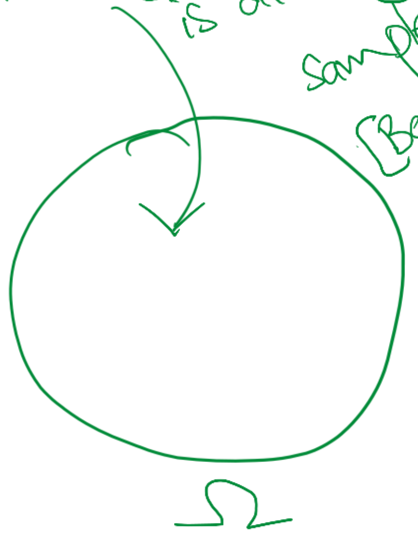


Probability – Axioms

There are three basic axioms of probability from which everything else can be derived

- assume we have a sample space Ω and events $A, B \subseteq \Omega$
- the probability of the event A is denoted $\mathbf{P}[A]$
- events A and B are mutually exclusive if $A \cap B = \emptyset$

Outcome of experiment is always from the sample space [Because that's how it is defined!]



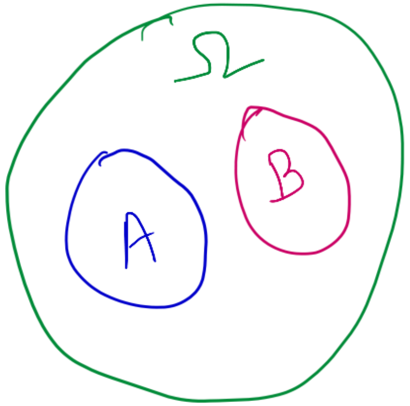
Axiom1: $\mathbf{0 \leq P[A] \leq 1}$ **for all events** $\mathbf{A \subseteq \Omega}$

Axiom2: $\mathbf{P[\Omega] = 1}$

Probability – Axioms

There are three basic axioms of probability from which everything else can be derived

- assume we have a sample space Ω and events $A, B \subseteq \Omega$
- the probability of the event A is denoted $\mathbf{P}[A]$
- events A and B are mutually exclusive if $A \cap B = \emptyset$



Axiom1: $0 \leq \mathbf{P}[A] \leq 1$ for all events $A \subseteq \Omega$

Axiom2: $\mathbf{P}[\Omega] = 1$

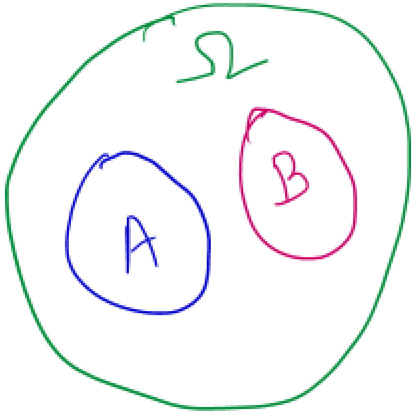
Axiom3: if A and B are mutually exclusive, then $\mathbf{P}[A \cup B] = \mathbf{P}[A] + \mathbf{P}[B]$

(This is essentially the sum rule, manifested in Probability theory)

Probability –
Two Dice Example Revisited In view of Axiom # 3

Axiom3: if A and B are mutually exclusive, then $P[A \cup B]=P[A]+P[B]$

(This is essentially the sum rule, manifested in Probability theory)



Rolling two dice

- sample space $\Omega = \{ (i,j) \mid 1 \le i \le 6 \wedge 1 \le j \le 6 \}$
  - $6 \cdot 6 = 36$ possible outcomes (using the product rule) and all equally likely
- probability that sum is even (*using Axiom #3*):

$$P[\text{both even}] + P[\text{both odd}]$$

$$= \frac{\text{ways first dice can be even} \times \text{ways second dice can be even}}{\text{Total number of ways two dice can be thrown}} + \frac{\text{ways first dice can be odd} \times \text{ways second dice can be odd}}{\text{Total number of ways two dice can be thrown}}$$

$$= (3 \cdot 3)/36 + (3 \cdot 3)/36$$

$$= 1/2$$

Probability – Properties from Axioms

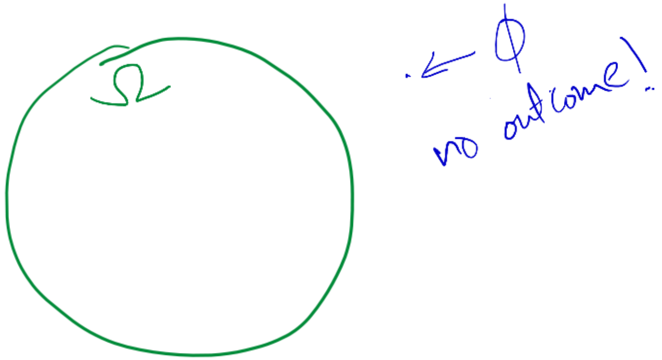
Axiom1: $0 \leq P[A] \leq 1$ for all events $A \subseteq \Omega$

Axiom2: $P[\Omega]=1$

Axiom3: if A and B are mutually exclusive, then $P[A \cup B]=P[A]+P[B]$

Some important properties

- $P[\emptyset]=0$



Probability – Properties from Axioms

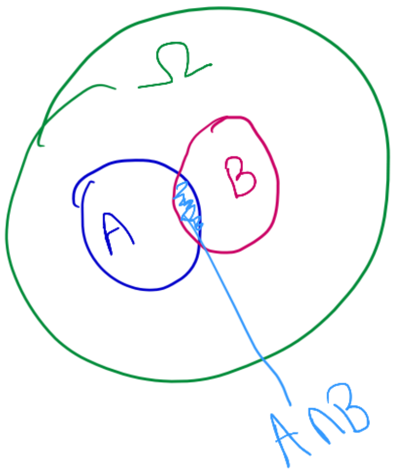
Axiom1: $0 \leq P[A] \leq 1$ for all events $A \subseteq \Omega$

Axiom2: $P[\Omega]=1$

Axiom3: if A and B are mutually exclusive, then $P[A \cup B]=P[A]+P[B]$

Some important properties

- $P[\emptyset]=0$
- $P[A \cup B]=P[A] + P[B] - P[A \cap B]$ for any events A and B



This is essentially the "inclusion-exclusion" principle of counting.

Probability – Properties from Axioms

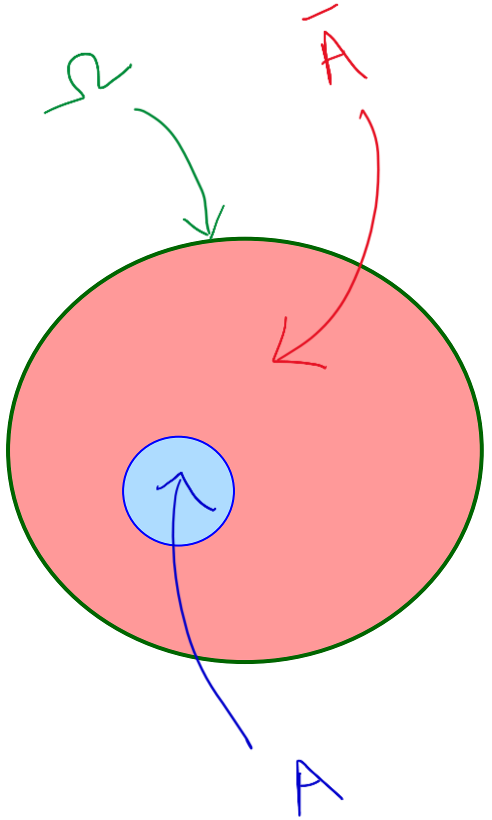
Axiom1: $0 \le P[A] \le 1$ for all events $A \subseteq \Omega$

Axiom2: $P[\Omega]=1$

Axiom3: if A and B are mutually exclusive, then $P[A \cup B]=P[A]+P[B]$

Some important properties

- $P[\emptyset]=0$
- $P[A \cup B]=P[A] + P[B] - P[A \cap B]$ for any events **A** and **B**
- $P[A]=1-P[\bar{A}]$ for any event **A** ($\bar{A}$ is complement of A)



A Quick Overview of Probability – Outline

- Probability – Introduction & Definitions
- Axioms and Properties of Probability
- Conditional Probability

Conditional probabilities

- Probability of an event, <u>given that</u> another event happened.

- E.g. given than that a dice throw is even, what is the probability that it is a 4?

Conditional probabilities

For events A and B, if $P[B]>0$, then the conditional probability of A given B is defined by: $P[A|B]$

Conditional probabilities

For events A and B, if $P[B]>0$, then the conditional probability of A given B is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

Conditional probabilities

For events A and B, if P[B]>0, then the conditional probability of A given B is defined by:

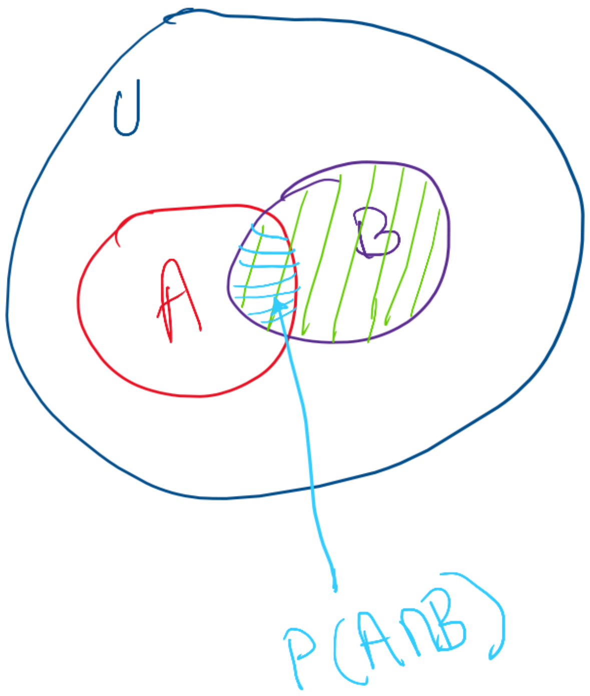
$$P[A|B] = \frac{P[A\cap B]}{P[B]}$$

Think of it this way:

Given that you have been told that event <u>B</u> has already happened, THAT is now your *Universe,* and the size of set B (its cardinality) will now be in the denominator (rather than the size of the original Universe).

For the numerator, you need to count the number of events of interest *inside your new Universe*. These are those elements of A (the event of interest) that are *also* in B (which is now the new Universe). That is, $|A\cap B|$.

So:
$P[A|B] = |A\cap B| \;/\; |B|$, or equivalently
$P[A|B] = P[A\cap B] \;/\; P[B]$



$$P[A|B] = \frac{P(A\cap B)}{P(B)}$$

Conditional probabilities

For events A and B, if $P[B]>0$, then the conditional probability of A given B is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

- need $P[B]>0$ as otherwise we are dividing by zero

Conditional probabilities

For events **A** and **B**, if **P[B]>0**, then the conditional probability of **A** given **B** is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

- need $P[B]>0$ as otherwise we are dividing by zero

If **A** and **B** are *independent** then $P[A|B] = P[A]$

**Independent events* means the occurrence of one event has no impact on the other. That is the same as saying that the Probability of A is the same whether or not B occurs, i.e. $P[A|B] = P[A]$.
E.g. given two dices, A is throwing first dice even, and B is throwing the second dice odd.
One has no impact on the other.

Note that this is entirely different what are called ***mutually exclusive** events*,
where the occurrence of one *precludes* the occurrence of the other.

E.g. A is throwing a dice even, and B is throwing a (the same) dice odd.

Conditional probabilities: Example

For events A and B, if P[B]>0, then the conditional probability of A given B is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

Example: given than that a dice throw is even, what is the probability that it is a 4? 

Conditional probabilities: Example

For events **A** and **B**, if **P[B]>0**, then the conditional probability of **A** given **B** is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

Example: given than that a dice throw is even, what is the probability that it is a 4?

– A the value of the die equals 4 (1 in 6 chance so probability 1/6)

Conditional probabilities: Example

For events A and B, if $P[B]>0$, then the conditional probability of A given B is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

Example: given than that a dice throw is even, what is the probability that it is a 4?

– A the value of the die equals 4 (1 in 6 chance so probability 1/6)
– B the value of the die is even (3 in 6 chance so probability 1/2)

Conditional probabilities: Example

For events **A** and **B**, if **P[B]>0**, then the conditional probability of **A** given **B** is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

Example: given than that a dice throw is even, what is the probability that it is a 4?

- A the value of the die equals 4 (1 in 6 chance so probability 1/6)
- B the value of the die is even (3 in 6 chance so probability 1/2)
- A∩B the value of the die is 4 and even (which is the same as A)

Conditional probabilities

For events A and B, if P[B]>0, then the conditional probability of A given B is defined by:

$$\mathbf{P}[A|B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]}$$

- need **P**[B]>0 as otherwise we are dividing by zero

Example rolling a fair die once

- A the value of the die equals 4 (1 in 6 chance so probability 1/6)
- B the value of the die is even (3 in 6 chance so probability 1/2)
- A∩B the value of the die is 4 and even (which is the same as A)
- $\mathbf{P}[A|B] = \mathbf{P}[A \cap B]/\mathbf{P}[B]$

Conditional probabilities

For events A and B, if P[B]>0, then the conditional probability of A given B is defined by:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

– need P[B]>0 as otherwise we are dividing by zero

Example rolling a fair die once

– A the value of the die equals 4 (1 in 6 chance so probability 1/6)
– B the value of the die is even (3 in 6 chance so probability 1/2)
– A∩B the value of the die is 4 and even (which is the same as A)
– $P[A|B] = P[A \cap B]/P[B] = (1/6)/(1/2) = 1/3$

Conditional probabilities

For events A and B, if $\mathbf{P}[B]>0$, then the conditional probability of A given B is defined by:

$$\mathbf{P}[A|B] = \frac{\mathbf{P}[A \cap B]}{\mathbf{P}[B]}$$

– need $\mathbf{P}[B]>0$ as otherwise we are dividing by zero

Example rolling a fair die once

– A the value of the die equals 4 (1 in 6 chance so probability 1/6)
– B the value of the die is even (3 in 6 chance so probability 1/2)
– $A \cap B$ the value of the die is 4 and even (which is the same as A)
– $\mathbf{P}[A|B] = \mathbf{P}[A \cap B]/\mathbf{P}[B] = (1/6)/(1/2) = 1/3$
  - $A|B$ means value is 4 given the value is even (1 in 3 chance, so probability 1/3)

Bayes' Rule

Motivation

Conditional Probability:

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

$$P[A|B] = \frac{P(A \cap B)}{P(B)}$$

In some cases, we don't know $P[A \cap B]$

- that is, we do not have enough information about overlap between A and B

But we are often in the situation that we *can* compute $P[B|A]$

- in general the two values $P[A|B]$ and $P[B|A]$ can be completely different

Typically, this type of problem occurs where we:

- want to know the probability of some event given some evidence
  - the likelihood I have a disease given that my blood test was positive
- but we only know the probability of the evidence given the event
  - if you have the disease, the blood test is positive 95% of the time

Motivating problem

Assume that

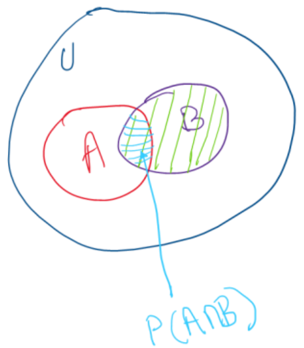
- 5% of incoming emails are spam (probability an email is spam: `0.05`)
- 50% of spam emails contain the word `'beneficiary'` (probability: `0.5`)
- 6% of all emails contain the word `'beneficiary'` (probability: `0.06`)

What is the probability that an email is spam, given that it contains the word `'beneficiary'`?

The Baye's Rule

- Starting point: Conditional Probability

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$



$$P[A|B] = \frac{P(A \cap B)}{P(B)}$$

Back to: Motivating problem

Assume that

- 5% of incoming emails are spam (probability an email is spam: `0.05`)
- 50% of spam emails contain the word `'beneficiary'` (probability: `0.5`)
- 6% of all emails contain the word `'beneficiary'` (probability: `0.06`)

What is the probability that an email is spam, given that it contains the word `'beneficiary'`?

The Baye's Rule

- Starting point: Conditional Probability

$$P[A|B] = \frac{P[A \cap B]}{P[B]}$$

$$P[A|B] = \frac{P(A \cap B)}{P(B)}$$

- P[A∩B] not known
- But: P[B|A] *is* known
- Express P[A∩B] in terms of P[B|A]?
  - $P[B|A] = \frac{P[B \cap A]}{P[A]}$
  - $P[B \cap A] = P[B|A].P[A]$
  - Since $P[B \cap A] = P[A \cap B]$
  - $P[A \cap B] = P[B|A].P[A]$ → replace in the conditional probability equation for P[A|B]

$$\boxed{P[A|B] = \frac{P[B|A]P[A]}{P[B]}}$$

Probability – Bayes' rule

Bayes' rule gives a consistent rule to take some prior belief and combine it with observed data to make a new prediction

We often phrase this as some **hypothesis H** we want to know, given some **data D** we observe, and we can write Bayes' Rule as:

$$\boxed{P[H|D] = \frac{P[D|H]P[H]}{P[D]}}$$

in other words

- to work out how likely a hypothesis is **true** given some data
- but only know how likely the data is if the hypothesis was **true**
- we can use Bayes' rule to solve the problem

Hypothesis and Data

Assume that

- 5% of incoming emails are spam (probability an email is spam: `0.05`)
- 50% of spam emails contain the word `'beneficiary'` (probability: `0.5`)
- 6% of all emails contain the word `'beneficiary'` (probability: `0.06`)

What is the probability that an email is spam, given that it contains the word `'beneficiary`'?

Hypothesis (H): Email is Spam

Data (D): Email has the word "beneficiary"

To find: P[H|D] = Probability that the hypothesis (email is spam) is true, given the Data (email contains the word beneficiary)

When what we know is: P[D|H] = Probability of finding data D (email contains the word beneficiary), when the hypothesis is true (email is email)

Bayesian Inference

- Predicting (Inferring) the future, given some data, by applying Baye's rule

- A very active area of investigation and application

- "Competes" with deep learning
  - Though often complementary

Probability – Problem set 3.1 – hints

Problem 3 hint:

- Let A be the event that the string contain at least two consecutive 0’s
- Let B be the event that the first bit is a 1
- Use conditional probability – what do you need to compute, $P[A|B]$ or $P[B|A]$?

Problem 4 hint:

- Let F be the event corresponding to dangerous fires taking place
- Let S be the event of smoke being produced
- Use Bayes’ law – what do you need to compute, $P[F|S]$ or $P[S|F]$?

Probability – Bayes' rule – Problem set 3.1 – problem 5

Suppose there are two boxes of balls

- the first box contains 2 white balls and 3 blue balls
- the second box contains 4 white and 1 blue ball.

Suppose you choose a box at random and then select a ball from that box at random

What is the probability that a ball from the first box was chosen, given you selected a blue ball?

- Let A_i be the event: choose the i^{th} box. Clearly $P[A_1] = 1/2$ and $P[A_2] = 1/2$
- Let B be the event: a blue ball is chosen. Then $P[B|A_1] = 3/5$ and $P[B|A_2] = 1/5$
- We want to find $P[A_1 \mid B]$
- Since $P[A_1] + P[A_2] = 1$, we can use Bayes' law:

$$\boxed{P[A_1|B] = \frac{P[B|A_1]P[A_1]}{P[B|A_1]P[A_1] + P[B|A_2]P[A_2]} = \frac{3/10}{3/10 + 1/10}}$$

Probability – Problem set 3.1 – problem 6 hint

A box contains 3 yellow balls and 5 red balls.

A ball is chosen at random from the box, then replaced in the box along with two other balls of the same colour.

- Hint: let R_i be the event the i^{th} ball is red and Y_i the i^{th} ball is yellow

- If a second ball is now chosen at random from the box, what is the probability that it will be red?
  - compute $P[R_2]$ knowing that the second ball is dependent on the choice of the first
- Given that the second ball is red, what is the probability that the first ball was yellow?
  - compute $P[Y_1 \mid R_2]$ using Bayes' law

Bayesian Inference

Predicting (Inferring) the future, given some data, by applying Bayes' rule

A very active area of investigation and application

"Competes" with deep learning

- though often complementary

Counting and Probability

- Product rule ✅
- Sum rule and the inclusion–exclusion principle ✅
- Combining product rule and sum rule ✅
- Pigeon–hole principle ✅
- Permutations and combinations ✅

- Probability – introduction and definitions ✅
- Axioms and properties of probability ✅
- Conditional probability and Bayes' rule ✅