Classical probability

Sample space Ω consists of a finite number of outcomes

$$\Omega = \{o_1, o_2, \cdots, o_n\}$$

Each outcome is equally likely to happen. Since we should have $P(\Omega) = 1$, this yields

$$P(o_i) = \frac{1}{n}, \, i = 1, 2, \ldots, n$$

The probability of an event E that represents k outcomes is then defined as the ratio of the number of outcomes favourable to E to n, that is

$$P(E) = \frac{k}{n}$$

Classical probability is a probability

1. $0 \le \dfrac{k}{n} \le 1$ for $k, n \in \mathbf{Z}^{+}, k \le n$

2. $P(\Omega) = \dfrac{n}{n} = 1$

3. $A \cap B = \varnothing \Rightarrow |A \cup B| = |A| + |B|$

Problems to solve

- what is the probability of a full house in poker?

- what are the chances of winning the second prize in a six-out-of-forty-nine lottery game?

- if a word is written at random composed of six letters, what are the chances that it will contain at least one "a"?

- in a random sample of n people, what is the probability that at least two of them will have the same birthday? (calculate first with leap years neglected and then included)

- n letters are written to different adressees and n envelopes provided with addresses, letters are then inserted at random into the envelopes. What are the chances that no addressee will get the right letter?

Discrete probability

This type of probability can be defined if the sample space is composed of at most countable number of outcomes

$$\Omega = \{o_1, o_2, \ldots, o_n\} \text{ or } \Omega = \{o_1, o_2, o_3, \ldots\}$$

finite set

infinite countable set

Each o_i is then assigned a non-negative real number p_i such that

$$\sum p_i = 1$$

We put

$$P(E) = \sum_{o_i \in E} p_i$$
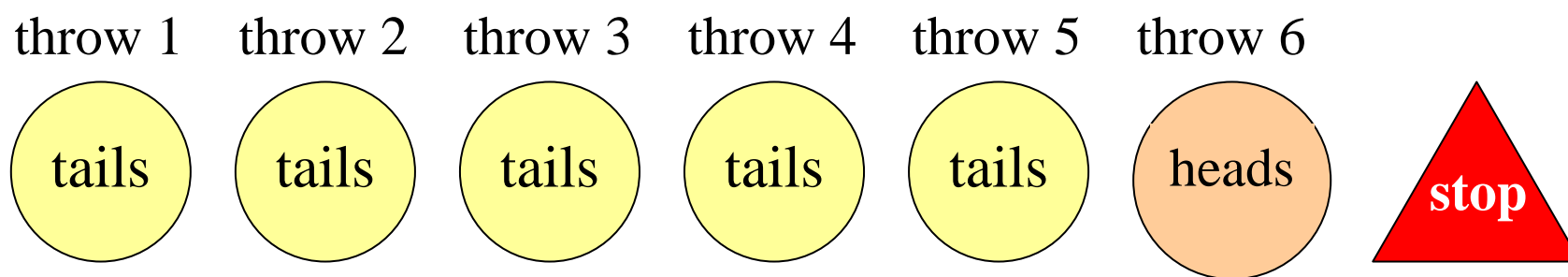
Discrete probability is a probability

Properties 1 and 2 follow from the fact that $\sum p_i = 1$

To prove property 3, we again use the implication

$$A \cap B = \varnothing \Rightarrow |A \cup B| = |A| + |B|$$

Example

An experiment is designed as tossing a fair coin repeatedly until "heads" appears for the first time. All possible outcomes of such an experiment are formed by sequences of tosses each with a series of tails finished up with heads. Obviously, each such outcome can be represented by a number n denoting the number of heads in the sequence.

throw 1	throw 2	throw 3	throw 4	throw 5	throw 6
tails	tails	tails	tails	tails	heads	stop

$$n = 5$$

Clearly, there are 2^n sequences of length n while only one of them represents the outcome n, which is $n+1$ tosses long, and so it is natural to put

$$p_n = \frac{1}{2^{n+1}}$$

We have

$$\sum_{i=0}^{n} p_i = \sum_{i=0}^{n} \frac{1}{2^{i+1}} = \sum_{i=1}^{n} \frac{1}{2^i} = \frac{1}{2} + \frac{1}{4} + \cdots = 1$$

and so we have defined a discrete probability, which can be used to calculate the probability of events.

For example, the chances of heads appearing no sooner than after 2 tosses are
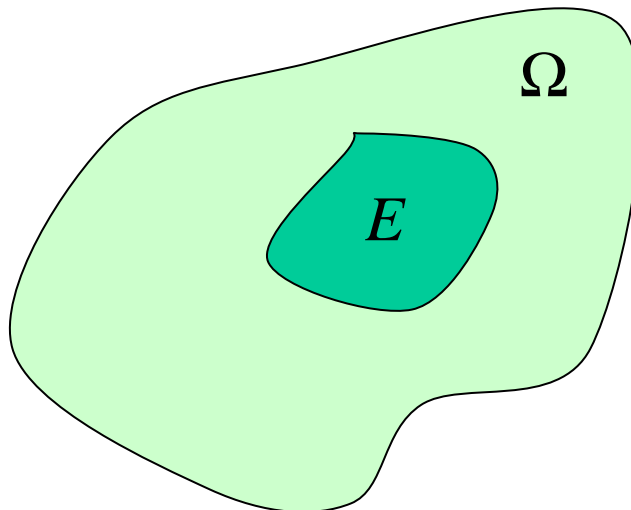
$$\frac{1}{16} + \frac{1}{32} + \cdots = 1 - \frac{1}{2} - \frac{1}{4} - \frac{1}{8} = \frac{1}{8}$$

Geometrical probability

If the sample space Ω can be represented by an area on a straight line, in a plane or in space and events by measurable subsets, then for an event E we can put

$$P(E) = \frac{m(E)}{m(\Omega)}$$

where $m(E)$ and $m(\Omega)$ are set measures (length, area, volume)

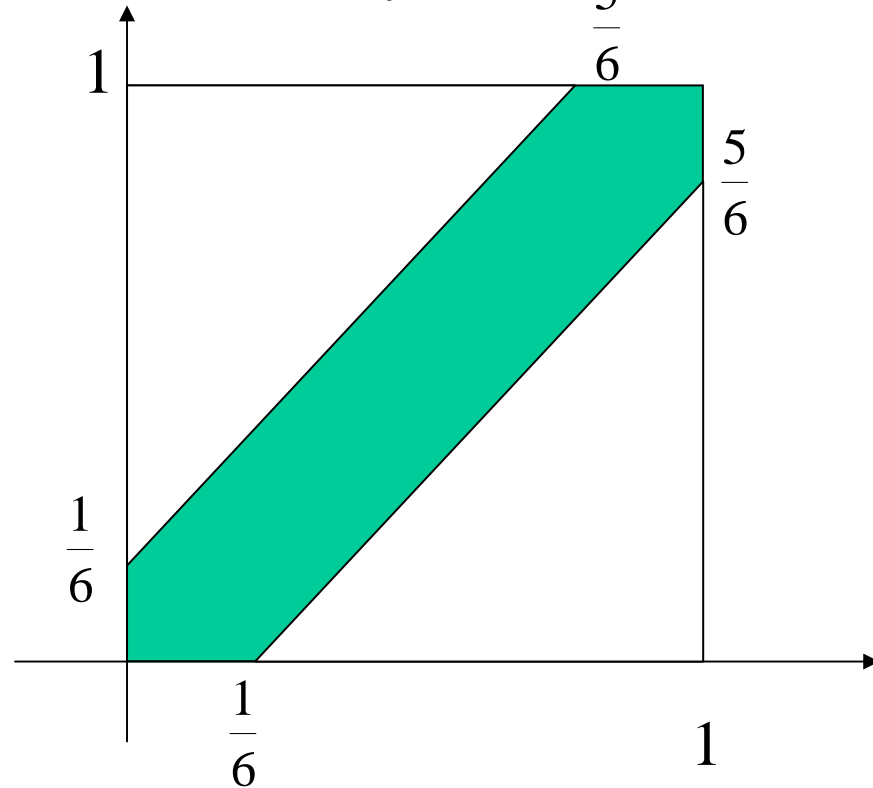Geometrical probability is a probability.

Properties 1 and 2 are obvious.

Property 3 follows from the axioms of the theory of measure.

Example

Two persons *A* and *B* agree to meet in a place between 1 pm and 2 pm. However, they don't specify a more precise time. Each of them will wait for ten minutes and leave if the other doesn't turn up. What are the chances that *A* and *B* really meet ?
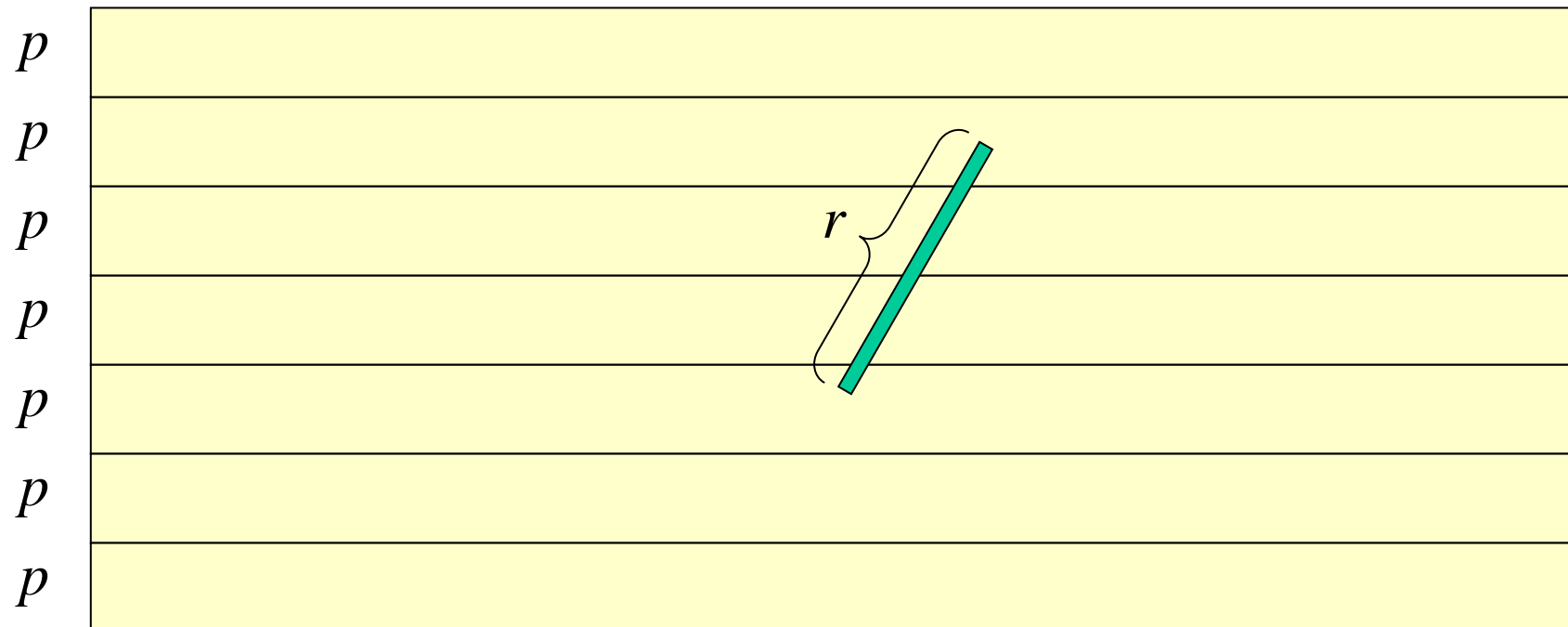
$$\Omega = \left\{ (t_A, t_B) \mid t_A, t_B \in [0,1] \right\}$$

$$E = \left\{ (t_A, t_B) \mid |t_A - t_B| \leq \frac{1}{6} \right\}$$

Problem to solve

We throw a rod of length r cm at random on the floor. The floor is made from planks each p cm wide. What is the probability that the rod lying on the floor will intersect at least one gap between the planks ?

$$\Omega$$

σ-fields of events

The above example shows that, unlike the finite sample space, not every subset of S is an event. We were only able to deal with those subsets that could be measured. In the theory of measure, only those systems S of subsets of a universal set S are considered that have the following properties:

(i) $\Omega \in \mathsf{S}$

(ii) $E \in \mathsf{S} \Rightarrow \overline{E} \in \mathsf{S}$

(iii) $E_1, E_2, E_3, \ldots \in \mathsf{S} \Rightarrow E_1 \cup E_2 \cup E_3 \cup \ldots \in \mathsf{S}$

In the last property, $E_1, E_2, E_3, \ldots$ denotes either a finite or an infinite countable sequence of subsets.

Such systems are called σ-fields.

$$\Omega$$

Probabilistic space

In subsequent considerations we will always assume that a sample space Ω, a σ-field S of events are given with a probability P mapping S into (0,1).

This triple is sometimes called a probabilistic space

$$P = (\Omega , S , P)$$

Conditional probability

What are the chances that I receive a straight flush (event *S*) ?

In each suite there are 10 straight flushes, the lowest beginning with an ace and the highest with a ten so using the classical probability, we have
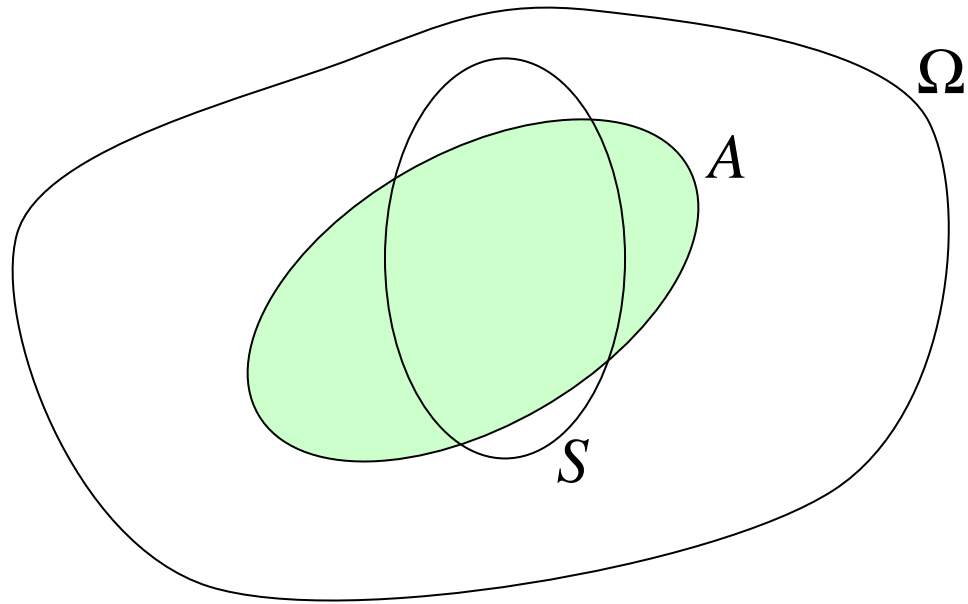
$$P(S) = \frac{(4)(10)}{C_{52}^5}$$

How would I determine the probability of such an event now that the first card has been dealt and I see that it is an ace (event *A*) ?

There are exactly eight straight flushes (two in each suit) with an ace and so

$$P(S \mid A) = \frac{8}{C_{52}^5 - C_{48}^5}$$

Calculating $P(S)$ and $P(S|A)$ we see that $P(S)$ is about 1.7 times higher that $P(S|A)$.

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

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

$P(S \mid A)$ is actually the probability of S as it changes if we receive additional information from the fact that A has occurred

This provides motivation for the definition of conditional probability, we put

$$P(E \mid A) = \frac{P(E \cap A)}{P(A)} \qquad P(A) \neq 0$$

reading $P(E \mid A)$ is the conditional probability of S given A.

We can also write

$$P(A \cap E) = P(A)P(E \mid A)$$

If $P(A \mid E) = P(A)$, we can show that also $P(E \mid A) = P(E)$

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

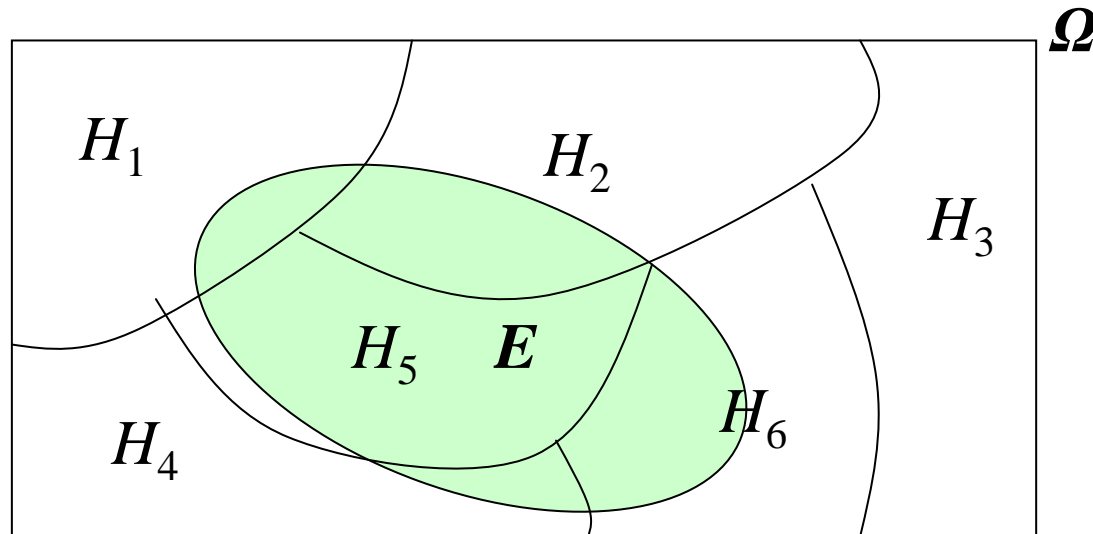If $P(A \mid E) = P(A)$ and $P(E \mid A) = P(E)$ we say that

_A_ and _E_ are independent

A and E are independent whenever $P(A \cap E) = P(A)P(E)$

The law of total probability

Conditional probability can be used to determine the probability of "sophisticated" events. The idea is to divide all the events into several categories. These categories are also called hypotheses. The sample space is partitioned using hypotheses H_0, H_1, ..., H_n, that is,

$$\Omega = H_1 \cup H_2 \cup \ldots H_n, \; H_i \cap H_j = \varnothing, \; i \neq j$$

We can write $E = (E \cap H_1) \cup (E \cap H_2) \cup ... \cup (E \cap H_n)$ with

$(E \cap H_i) \cap (E \cap H_j) = \varnothing, i \neq j$
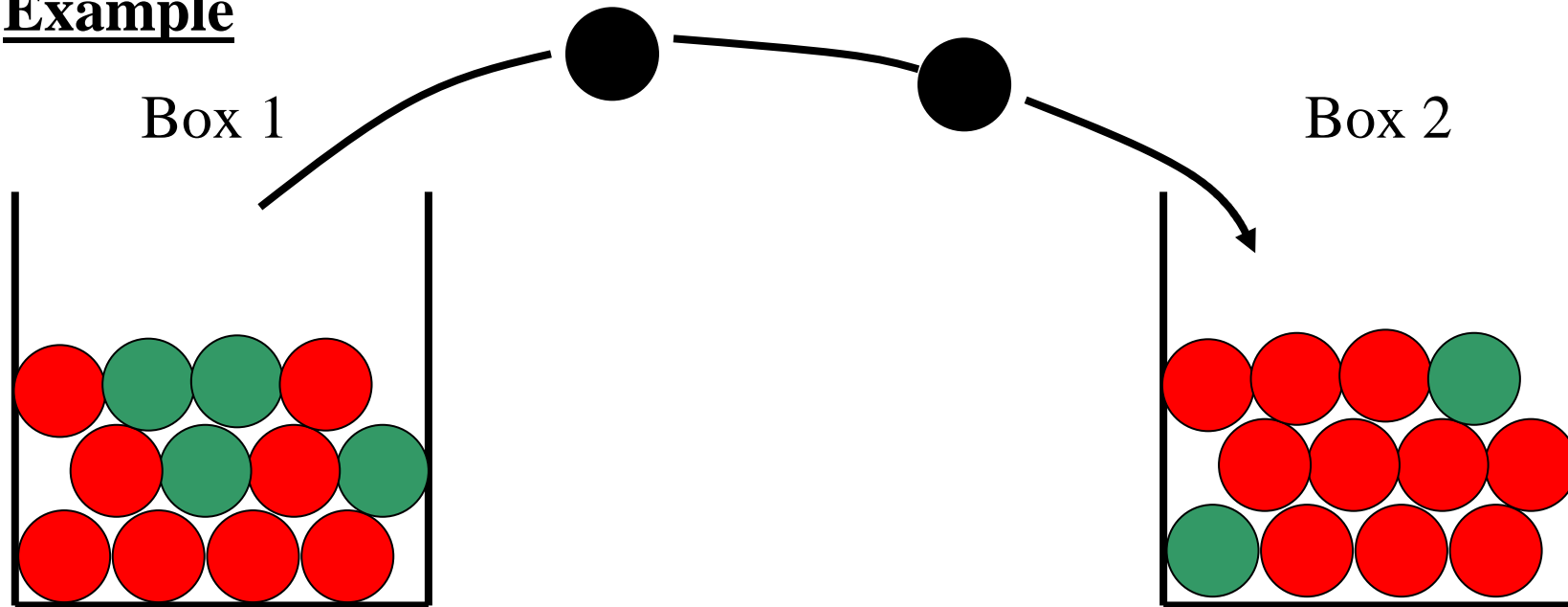
This means that

$P(E) = P(E \cap H_1) + P(E \cap H_2) + ... + (E \cap H_n)$

which yields

$$P(E) = P(E \mid H_1)P(H_1) + P(E \mid H_2)P(H_2) + \cdots + P(E \mid H_n)P(H_n)$$

This formula is called the law of total probability.

Example

Box 1

Box 2

8 red balls, 4 green balls

10 red balls, 2 green balls

Two balls are moved at random from box 1 to box 2. What are the chances that a ball subsequently drawn at random from box 2 will be red ?

In this experiment we will differentiate six possible outcomes:

$$\Omega = \{(0,R),(1,R),(2,R),(0,G),(1,G),(2,G)\}$$

where, for example, $(1,R)$ means that of the two balls moved

one was red and red ball was subsequently drawn from box 2.

We put $H_i = \{(i,R),(i,G)\}, i = 0,1,2$

H_0, H_1, H_2 obviously fulfill the conditions of hypotheses and so if we denote by R the event that a red ball was drawn from box 2, we can write

$$P(R) = P(R\,|\,H_0)P(H_0) + P(R\,|\,H_1)P(H_1) + P(R\,|\,H_2)P(H_2)$$

Then we can calculate

$$P(H_0) = \frac{C_4^2}{C_{12}^2}, \; P(H_1) = \frac{C_4^1 C_8^1}{C_{12}^2}, \; P(H_2) = \frac{C_8^2}{C_{12}^2}$$

$$P(R \mid H_0) = \frac{10}{14}, \; P(R \mid H_1) = \frac{11}{14}, \; P(R \mid H_2) = \frac{12}{14}$$

So that $P(R) = \dfrac{10}{14} \cdot \dfrac{6}{66} + \dfrac{11}{14} \cdot \dfrac{32}{66} + \dfrac{12}{14} \cdot \dfrac{28}{66} \approx 0.80952$

Example

Suppose Peter and Paul initially have *m* and *n* dollars, respectively. A ball, which is red with probability *p* and black with probability $q = 1 - p$, is drawn from an urn. If a red ball is drawn, Paul must pay Peter one dollar, while Peter must pay Paul one dollar if the ball drawn is black. The ball is replaced, and the game continues until one of the players is ruined. Find the probability of Peter's ruin.

Let $Q(x)$ denote the probability of Peter's ruin if he has x dollars. In this situation there are two possibilities, which we will use as hypotheses W- Peter wins and L - Peter loses. Then, by the law of total probability, we can write

$$Q(x) = Q(x+1)P(W) + Q(x-1)P(L)$$

$$Q(x) = Q(x+1)p + Q(x-1)q$$

We want to determine $Q(m)$. The above equation can be solved as a difference equation with boundary conditions

$$Q(m+n) = 0 \text{ and } Q(0) = 1.$$

Result

$$Q(m) = p^{2m} q^m \frac{p^n - q^n}{p^{m+n} - q^{m+n}} \quad \text{if} \quad p \neq q$$

$$\text{and} \quad Q(m) = \frac{n}{m+n} \quad \text{if} \quad p = q$$

Law of inverse probability

Three different companies *A*, *B*, *C* supply eggs to a supermarket. Supplies from *A* account for one third of the eggs sold, *B* supplies a quarter of all eggs and *C* the rest. It is known that with *A* one egg in a hundred thousand will be decayed, with B this ration is 1:200000 and with C it is one egg in 500 thousand. With what probability an egg that you have bought and found it decayed has been supplied by *B*?

1/3	1/4	5/12
A	B	C
1:100 000	1:200 000	1:500 000

The event D will denote the event that the bought egg is decayed. A, B, and C will be hypotheses denoting the events that the bought egg comes from A, B, C respectively.

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

Using the law of total probability, we can further write

$$P(B \mid D) = \frac{P(D \mid B)P(B)}{P(D \mid A)P(A) + P(D \mid B)P(B) + P(D \mid C)P(C)}$$

$$P(B \mid D) = \frac{\dfrac{1}{4 \cdot 200\,000}}{\dfrac{1}{3 \cdot 100\,000} + \dfrac{1}{4 \cdot 200\,000} + \dfrac{5}{12 \cdot 500\,000}} = \frac{4}{15}$$

The above method can be used to derive such formulas in a general case. Let $H_1, H_2, \ldots, H_n$ be hypotheses and E an event. The probability of E happening because of H_i can be calculated using the following formulas

$$P(H_i \mid E) = \frac{P(E \mid H_i) P(H_i)}{\displaystyle\sum_{j=1}^{n} P(E \mid H_j) P(H_j)}, \quad i = 1, 2, \ldots, n$$

known as Bayes's theorem (after the 18th-century English clergyman Thomas Bayes) or the law of inverse probability.