BASICS OF PROBABILITY

How likely is it that the event A occurs? In everyday situations expressions like

impossible
unlikely
possible
conceivable
probable
likely
certain

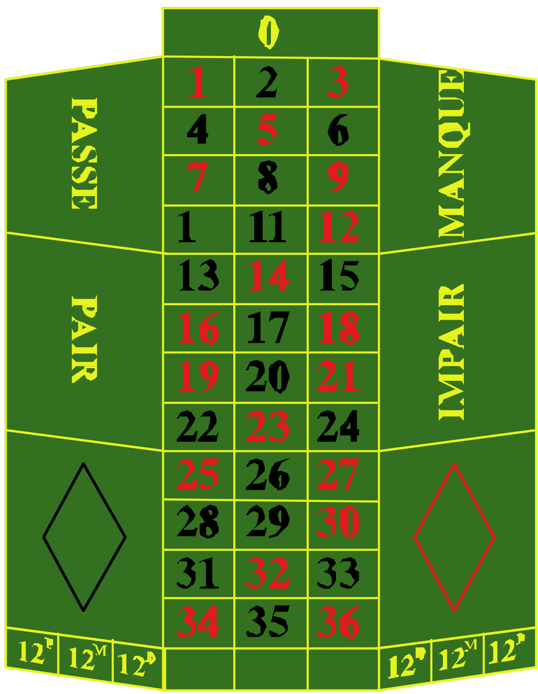
For mathematics this list is not accurate enough. There an event A is mapped to a number p(A) between 0 and 1 according to its likelihood.

In what follows crucial expressions and correlations are introduced with the help of roulette.
Roulette is a game of chance where the player has to guess the number of the compartment into which a little ball is going to fall. The compartments are numbered from 0 to 36 and arranged along the rim on top of a turning wheel. 18 of the numbers are red, 18 are black and the number 0 is green.
The player now places his stake on the tableau either on a number or on a combination of numbers.
If his number or his combination of numbers shows up he wins and he gets paid according to the formula:

$$\text{payment}(\text{placement}) = \frac{36}{n} \cdot \text{stake}$$

n = number of compartments forming the placement

If his number or his combination of numbers does not show up he loses his stake.



1 Fill in he following table.

putting stake s on	payed back
MANQUE/PASSE	
PAIR/IMPAIR	
ROUGE/NOIR	
12^D / 12^M / 12^P	
1st, 2nd, 3rd column	
single number	
square	

MANQUE	the numbers from 1 to 18.
PASSE	the numbers from 19 to 36.
PAIR	the even numbers.
IMPAIR	the odd numbers.
ROUGE	the red numbers.
NOIR	the black numbers.
12^D	the numbers from 1 to 12.
12^M	the numbers from 13 to 24.
12^P	the numbers from 25 to 36.
square	four numbers that join the same vertex.

DEFINITION
A **probability experiment** (in the following short **experiment**) is an experiment with outcomes that are in a clearly defined range but can not be foretold.

EXAMPLE
Rolling the ball at roulette.

DEFINITION
The **sample space** $\boldsymbol{\Omega}$ is the set of all possible outcomes of an experiment.

EXAMPLE
At roulette the sample space is $\Omega = \{0, 1, 2, \ldots, 36\}$.

DEFINITION
An **event** is a subset of the sample space.

EXAMPLE
Event $A = \text{red numbers} = \{1, 3, 5, 7, 9, 12, 14, \ldots, 36\}$.

Event $B = \text{2nd column} = \{2, 5, 8, \ldots, 35\}$.

Event $C = 12^P = \{25, 26, \ldots, 36\}$.

Event $D = \text{number } 14 = \{14\}$.

DEFINITION
If an experiment has the outcome a then **the event A occurs** exactly if $a \in A$.

EXAMPLE
If number 14 is the outcome of a round at roulette then the events A, B and D but not C have occurred.

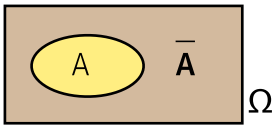
DEFINITION
A **probability function p** is a function that maps the set of all subsets of Ω to the interval $[0,1]$:

$$p: \quad \text{set of subsets of } \Omega \to [0,1]$$
$$A \mapsto p(A)$$

In addition the function p has to meet the following three requirements (Kolmogorov's axioms):

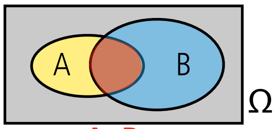
1) $p(A) \geq 0$
2) $p(\Omega) = 1$
3) If $A \cap B = \{\ \}$ then $p(A \cup B) = p(A) + p(B)$

REPETITION
In set theory there are signs used that need some explanation.



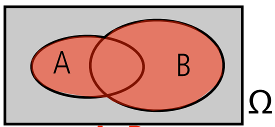
$a \in A$: a is an element of A

$\bar{A}$ = set that contains all the elements that do **not** belong to A
= complement of A



A∩B

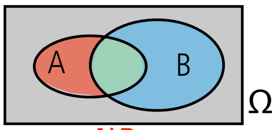
$|A|$ = number of elements in A



A∪B

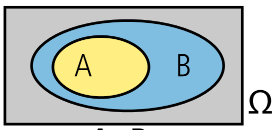
$A \cap B$ = set that contains all the elements of A and B that belong to both of them = intersection of A and B

$A \cup B$ = set that contains all the elements of A and B that belong to at least one of them = union of A and B



A\B

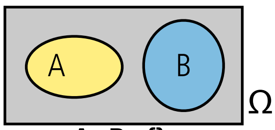
$A \setminus B$ = set that contains all the elements of A that do not belong to B = A without B



A⊂B

$A \subset B$: A is a subset of B

$\{\ \}$ = set with no elements = empty set



A∩B={}

$A \cap B = \{\ \}$: the intersection of A and B is the empty set = A and B are disjoint (have no element in common)

EXAMPLE

In roulette there is only a finite number of possible outcomes $\left(|\Omega| = 37\right)$ and are equally likely. The following definition of a probability function is quite obvious:

$$\begin{aligned} p: &\ \text{set of subsets of } \{0, 1, 2, \dots, 36\} \to [0,1] \\ &A \mapsto p(A) = \frac{|A|}{37} \end{aligned}$$

This definition meets all the requirements of a probability function:

1) $|A| \geq 0 \quad \Rightarrow \quad p(A) \geq 0$
2) $p(\Omega) = \frac{|37|}{37} = 1$
3) If $A \cap B = \{\ \}$ then $p(A \cup B) = \frac{|A \cup B|}{37} = \frac{|A| + |B|}{37} = p(A) + p(B)$

2 Calculate the probability of the events A, B, C and D from above:

a) $A = \text{red numbers} = \{1, 3, 5, 7, 9, 12, 14, \dots, 36\}$.

b) $B = \text{2nd column} = \{2, 5, 8, \dots, 35\}$.

c) $C = 12^P = \{25, 26, \dots, 36\}$.

d) $D = \text{number } 14 = \{14\}$.

This last example can be generalised.

DEFINITION

Experiments with a finite number of possible outcomes which are all equally likely are **Laplace experiments**.

THEOREM

With all Laplace experiments $p(A) = \frac{|A|}{|\Omega|}$ is the corresponding probability function.

Kolmogorov's axioms form the foundation of probability calculus. An axiom is a basic statement that is accepted as true. All theorems in probability calculus can be formally deducted from them:

1) $p(A) \geq 0$
2) $p(\Omega) = 1$
3) If $A \cap B = \{\ \}$ then $p(A \cup B) = p(A) + p(B)$

Andrey Kolmogorov (1903 - 1987), mathematician
His unwed mother died in childbirth and he was raised by his mother's sister at the estate of his grandfather, a wealthy nobleman. His father, an agriculturist, was exiled and killed during the Russian Revolution.
In 1919 his aunt adopted him and they moved to Moscow where he attended a gymnasium. In 1920 he started his studies of mathematics and already two years later published his first papers on set theory and differential calculus. He was the first to succeed in finding an axiomatic bases for probability calculus, one of the famous 23 unsolved problems presented by Hilbert in 1900.
He gained a high reputation not only for his scientific work but also for his dedication to teaching. He is said to be the most influential mathematician of the first half of the 20th century.

David Hilbert (1862 - 1943), mathematician
He was born in Königsberg and after graduating from the gymnasium he went to the University of Königsberg to study mathematics.
He received his doctorate when he was 23 years old and started lecturing immediately. In 1895 he was offered the chair of mathematics at the University of Göttingen which he accepted and where he stayed until he resigned in 1930.
He was the most influential and all-round mathematician of his time. He was the last one to oversee all the developments in all the different branches of mathematics.
For Hilbert mathematics was a complex game, absolutely independent from reality. His vision was the complete axiomatisation of mathematics, he did not strive for more and more knowledge but to put the collected knowledge on a secure basis.
In 1900 at the 2nd International Congress of Mathematics in Paris he delivered his famous speech in which he listed 23 then unsolved problems in mathematics. Some of them were solved in the years following, some have remained unsolved until today. But each of them proved to be of great significance for the development of mathematics.

From Kolmogorov's axioms the following theorem can be deducted:

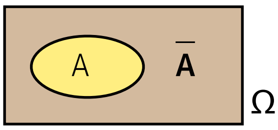
THEOREM

1) $p(\{\ \}) = 0$
2) $p(\overline{A}) = 1 - p(A)$
3) If $A \subset B$ then $p(A) \leq p(B)$
4) If $A = \{a_1, a_2, \dots, a_n\}$ then $p(A) = p(\{a_1\}) + p(\{a_2\}) + \dots + p(\{a_n\})$

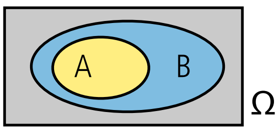
PROOF

1) Because of $\Omega \cap \{\ \} = \{\ \}$ the two sets are disjoint. From axiom 3 can be concluded that $p(\Omega \cup \{\ \}) = p(\Omega) + p(\{\ \})$.
   Because of $\Omega = \Omega \cup \{\ \}$ follows $p(\Omega) = p(\Omega) + p(\{\ \})$.
   With axiom 1 there is $1 = 1 + p(\{\ \})$ and therefore $p(\{\ \}) = 0$.

2) $\Omega = A \cup \overline{A}$, where A and $\overline{A}$ are disjoint. From axiom 3 can be concluded that $1 = p(\Omega) = p(A \cup \overline{A}) = p(A) + p(\overline{A})$.
   Therefore $p(\overline{A}) = 1 - p(A)$



3) If $A \subset B$ then $B = A \cup (B \setminus A)$, where A and B\A are disjoint.
   From axiom 3 can be concluded
   $p(B) = p(A \cup (B \setminus A)) = p(A) + p(B \setminus A)$.
   From axiom 1 can be concluded that $p(B \setminus A) \geq 0$.
   So $p(B) \geq p(A)$.



4) If $A = \{a_1, a_2, \dots, a_n\}$ then $A = \{a_1\} \cup \{a_2\} \cup \dots \cup \{a_n\}$, where all sets $\{a_1\}, \{a_2\}, \dots, \{a_n\}$ are disjoint. From axiom 3 can be concluded that $p(A) = p(\{a_1\}) + p(\{a_2\}) + \dots + p(\{a_n\})$.

NB
Theorem 4) states that if you know the probabilities of all possible outcomes of an experiment, i.e. $p(\{a\})$ for all $a \in \Omega$, the probabilities of any event A can be calculated.

DEFINITION
Be p a probability function. Then

$$d: \quad \Omega \to [0,1]$$
$$a \mapsto d(a) = p(a)$$

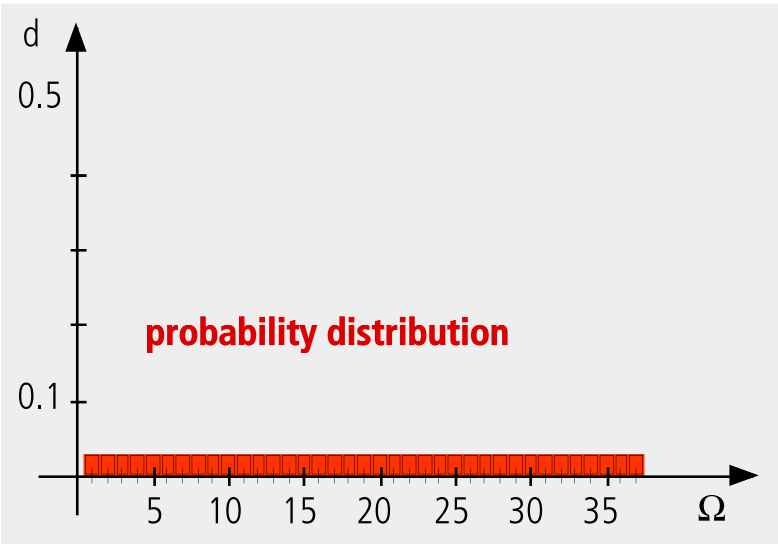
is the corresponding **probability distribution**.

EXAMPLE
In roulette

$$p: \text{ set of subsets of } \{0, 1, 2, \dots, 36\} \to [0,1]$$
$$A \mapsto p(A) = \frac{|A|}{37}$$

is the probability function. Its probability distribution is

$$d: \quad \{0, 1, 2, \dots, 36\} \to [0,1]$$
$$x \mapsto d(x) = \frac{1}{37}$$



Instead of single dots (d has no value between the integers in Ω) often a rectangle with the height d(a) and the width 1 is plotted. So not only the height but also the area of the rectangle corresponds to the probability of a.

The graphs of probability distributions of Laplace experiments always form a straight line.

EXERCISE

Calculate the probability that at roulette in three rounds you obtain

a) event A = "all three colours".

b) event B = "exactly twice red".

SOLUTION

To find the correct solution it is often helpful to look at a wrong one first.

Wrong way:

Be r = red, b = black and g = green.

A possible outcome in three rounds is any combination of r, b and g:

$\Omega = \{$rrr, rrb, rrg, rbr, rbb, rbg, rgr, rgb, rgg, brr, brb, brg, bbr, bbb, bbg, bgr, bgb, bgg, grr, grb, grg, gbr, gbb, gbg, ggr, ggb, ggg$\}$ $\Rightarrow$ $|\Omega| = 27$

a) $A = \{$rbg, rgb, brg, bgr, grb, gbr$\}$ $\Rightarrow$ $|A| = 6$ $\Rightarrow$ $p(A) = \frac{6}{27} = \frac{2}{9}$

b) $B = \{$rrb, rrg, rbr, rgr, brr, grr$\}$ $\Rightarrow$ $|B| = 6$ $\Rightarrow$ $p(B) = \frac{6}{27} = \frac{2}{9}$

These results are wrong because all the combinations of r, b and g are not equally likely!

rrr for example is much more likely than ggg.

But if all the possible outcomes are not equally likely then the experiment is not a Laplace experiment and the formula $p(A) = \frac{|A|}{|\Omega|}$ can not be used.

1st correct way:

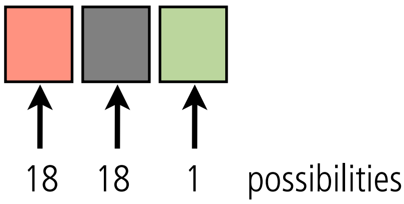
All possible outcomes of the experiment are the triples from $(0|0|0)$ up to $(36|36|36)$:

$$\Omega=\{(0|0|0),(0|0|1),\ldots,(36|36|35),(36|36|36)\}$$

They are all equally likely, so it is a Laplace experiment and the formula $p(A)=\frac{|A|}{|\Omega|}$ can be used. $\Rightarrow$ $|\Omega|=37^3=50{,}653$

a) $A=\{(1|2|0),(1|4|0),\ldots,(0|35|36)\}$

The big problem to solve is now the calculation of $|A|$. Let us assume that the first number is the red, the second the black one and the third the green one. How many possibilities are there to fill in these places?



So there are 18·18·1 = 324 possible triples, all different from each other.

These are only the triples with a red number at the first place, a black one at the second place and a green number at the end. But there might be triples with the black number at the first place, the red number at the second place and the green at the third place and so on.
So we can swap the places in the triples. The triple (**23**|**9**|**0**) for example can be altered into (**23**|**0**|**9**), (**9**|**23**|**0**), (**9**|**0**|**23**), (**0**|**23**|**9**), (**0**|**9**|**23**).
For each triple with the order red|black|green there are 6 possible triples with the same numbers but different positions.

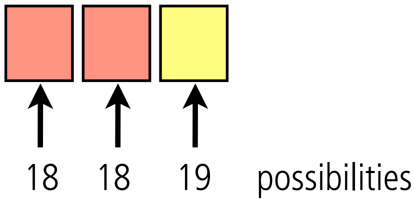
$\Rightarrow$ $|A|=324\cdot 6=1944$

$\Rightarrow$ $p(A)=\frac{|A|}{|\Omega|}=\frac{1{,}944}{50{,}653}\approx \mathbf{0.038}$

b) $B = \{(1|1|0), (1|1|2), \dots, (35|35|36)\}$

The evaluation of $|B|$ is even more complicated than the one of $|A|$.

Similar to before we assume that the first two positions are red and the third black or green. How many possibilities are there to fill in these places?



So there are $18 \cdot 18 \cdot 19 = 6156$ possible triples, all different from each other.

Again we can swap the positions in the triples.

Let us first look at triples with **different** red numbers:
The triple (**23**|**3**|**9**) for example can be altered into (**23**|**9**|**3**), (**9**|**23**|**3**), (**9**|**3**|**23**), (**3**|**23**|**9**) and (**3**|**9**|**23**). But the triple (**3**|**23**|**9**) can be altered into the same 6 triples! That leaves 3 possible triples for each triple with different red numbers.

Then we look at triples with **equal** red numbers:
The triple (**23**|**23**|**9**) for example can be changed only into (**23**|**9**|**23**) and (**9**|**23**|**23**).

For each triple with two red numbers there are 3 possible triples with the same numbers but different positions.

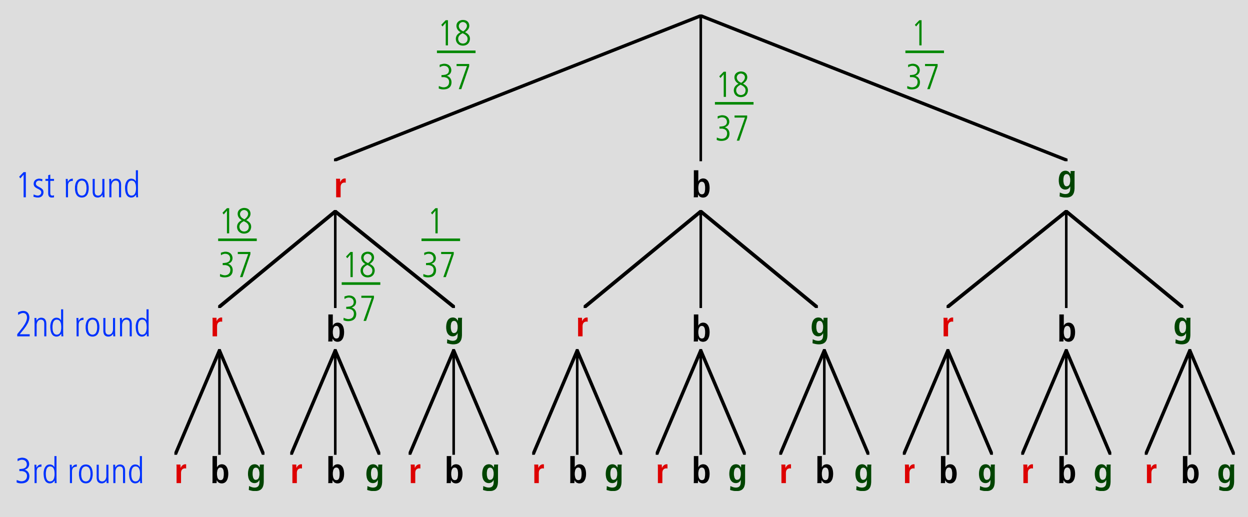
$\Rightarrow \quad |B| = 6156 \cdot 3 = 18,468$

$\Rightarrow \quad p(B) = \frac{|B|}{|\Omega|} = \frac{18,468}{50,653} \approx \mathbf{0.365}$

2nd correct way:

There is a much easier way than the first one.

The experiment is composed of three identical rounds at roulette. It can be visualised by a tree diagram:



On each level there are all possible outcomes of that particular part of the compound experiment.

Each branch from the top to the bottom corresponds to one possible outcome of the compound experiment.

The green numbers next to the branches indicate the probability that this branch is picked under the condition that the experiment is at that branching.

At each branching the corresponding part of the experiment is a Laplace experiment and the green numbers can be calculated with the formula $p(A)=\dfrac{|A|}{|\Omega|}$.

a) $A=\{rbg, rgb, brg, bgr, grb, gbr\}$

$$\Rightarrow \quad p(A)=\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{1}{37}+\frac{18}{37}\cdot\frac{1}{37}\cdot\frac{18}{37}+\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{1}{37}+\frac{18}{37}\cdot\frac{1}{37}\cdot\frac{18}{37}+\frac{1}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}+\frac{1}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}$$

$$=6\cdot\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{1}{37}\approx \mathbf{0.038}$$

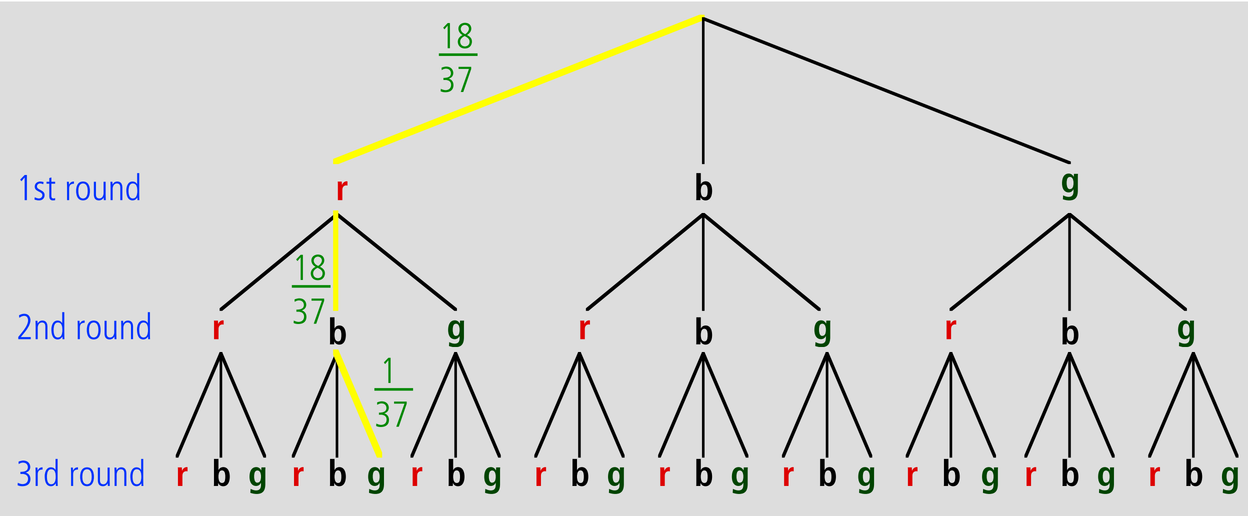
b) $B=\{rrb, rrg, rbr, rgr, brr, grr\}$

$$\Rightarrow \quad p(b)=\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}+\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{1}{37}+\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}+\frac{18}{37}\cdot\frac{1}{37}\cdot\frac{18}{37}+\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}+\frac{1}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}$$

$$=3\cdot\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{18}{37}+3\cdot\frac{18}{37}\cdot\frac{18}{37}\cdot\frac{1}{37}\approx \mathbf{0.365}$$

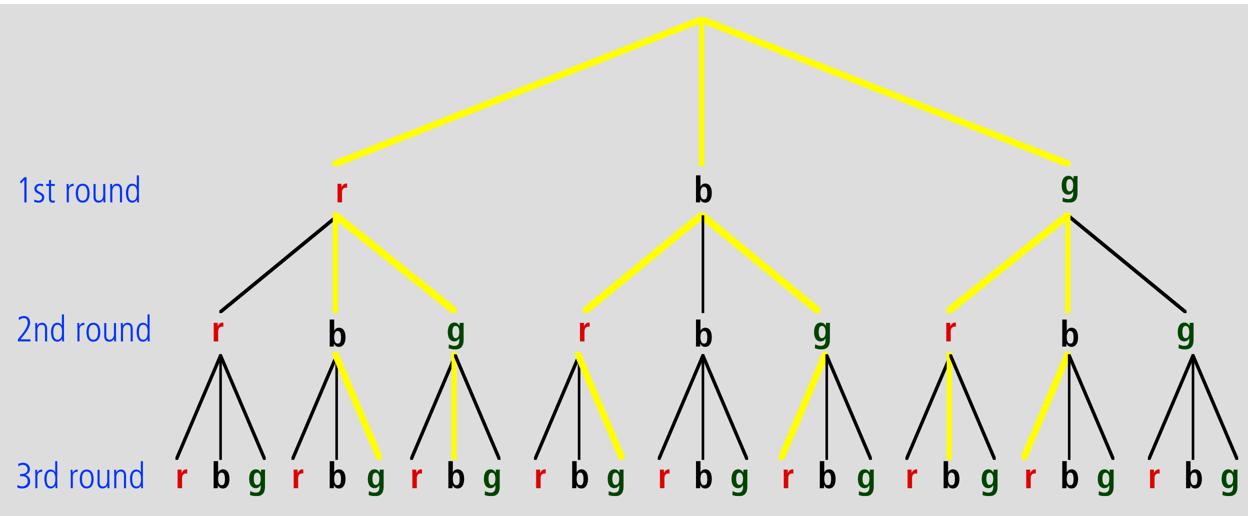
There are two rules to calculate probabilities with a tree diagram.

MULTIPLICATION RULE
The probability of a branch (from top to bottom) is the **product** of all green numbers along the branch.



$$p(rbg) = \frac{18}{37} \cdot \frac{18}{37} \cdot \frac{1}{37} = \frac{324}{50{,}653}$$

ADDITION RULE
The probability of an event is the **sum** of the probabilities of all branches forming the event.



$$p(\text{"all three colours"}) = 6 \cdot \frac{324}{50{,}653} = \frac{1{,}944}{50{,}653} \approx 0.038$$

3 At roulette, how many times do you have to put money on the same number to win at least once with a probability of at least 95 %?

4 At roulette, what is the probability that

a) five times in a row a red number shows up?

b) in the 5th round red shows up under the condition that in the four previous rounds there were four red numbers in a row?

5 Bart and Milhouse try out their catapult and shoot one pebble each at Principal Skinner's office window. Bart usually hits the target four times out of five, Milhouse is less skilful and hits it only seven times out of ten. What is the probability that the window is hit at least once?

6 Mr Bond claims to be able to distinguish a stirred from a shaken Martini by tasting only. His friends do not believe him and want to put him to the test. They set up a row of five identical glasses and fill them at random with either a stirred or a shaken Martini. Then they ask Mr Bond to taste them blindfolded and pick out the stirred ones. They will acknowledge his ability if he gets at least four glasses right.

a) What is the probability for Mr Bond to pass the test by just guessing?

b) What is the probability for Mr Bond to fail the test even if he can tell the glasses apart with a probability 90 %?

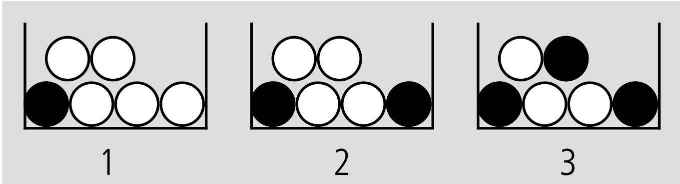
7 The owner of a fast-food place knows from experience that 80 % of his customers order kebab, 70 % want some extra ketchup and 60 % have a coke to wash it down. To reduce costs he decides to serve everybody kebab with ketchup and a coke without even asking.

a) What are the chances that a customer gets exactly what he wants?

b) What are the chances that the customer gets at least two things he wants?

8 The caliph of Baghdad had the habit of having thieves incarcerated and the right hand chopped off. But beforehand they got a chance to save the limb. They could choose blindfolded one of the three urns and draw a ball. If the ball was white they where freed immediately.

a) What is the probability to draw a white ball?



One day a clever thief asked whether he was allowed to move some balls - with eyes open - from one urn to the other. The caliph agreed assuming that this would not increase the thief's chance to escape unharmed.

b) Was the caliph right?

9 The "one-armed bandit" is a special kind of gambling machine in casinos. Basically there are three spinning drums with - in our case - ten different pictures on each of them. By pulling the lever (= arm) down the drums are set in motion and start spinning. They come to a random stop independently after a certain time and show a picture each.

a) You hit the jackpot if the three pictures are identical. What are the chances?

b) What are the chances that exactly two pictures are identical?

10 In a bag there are 26 wooden cubes with a different letter of the alphabet on each one. Now the bag is shaken to shuffle the cubes thoroughly. Then, one by one, four cubes are randomly taken out.

a) What are the chances that in the order they are taken out they form the word "MARS"?

b) What are the chances that the word "MARS"can be formed by them?

11 "Texas Hold'm" is a form of poker in which every player is given randomly five cards out of 52. In addition two more cards which belong to all players are openly placed on the table.

a) What are the chances for "All Fours" (four cards of the same rank, e.g. four nines)?

b) What are the chances for a "Full House" (three cards of the same rank and two cards of the same rank, e.g. three aces and two eights)?