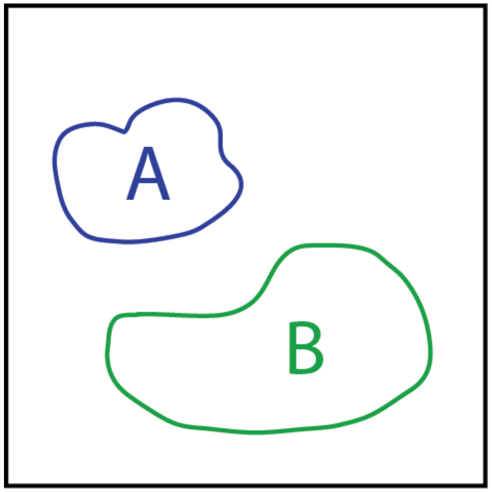
Probability

Chapter 5

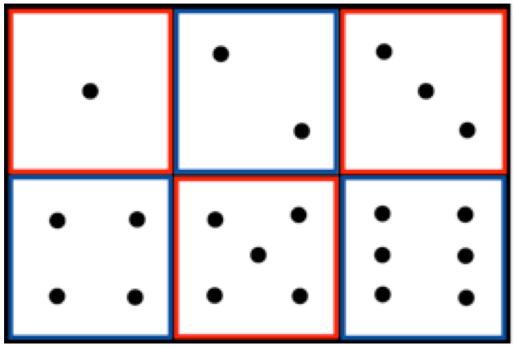
The probability of an event is its true relative frequency, the proportion of times the event would occur if we repeated the same process over and over again.

Two events are *mutually exclusive* if they cannot both be true.

A and B are mutually exclusive



Mutually exclusive



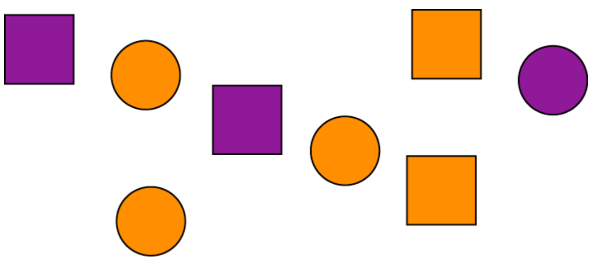
Mutually exclusive

Pr(*A* and *B*) = 0

Not mutually exclusive

Pr(*A* and *B*) ≠ 0

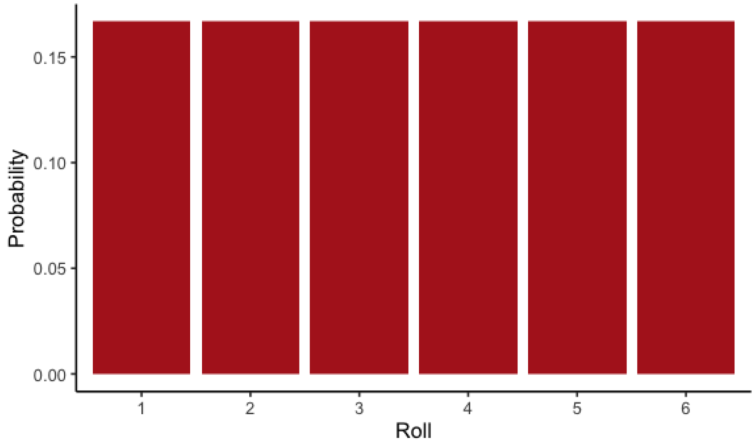
Pr(*purple* AND *square*) ≠ 0



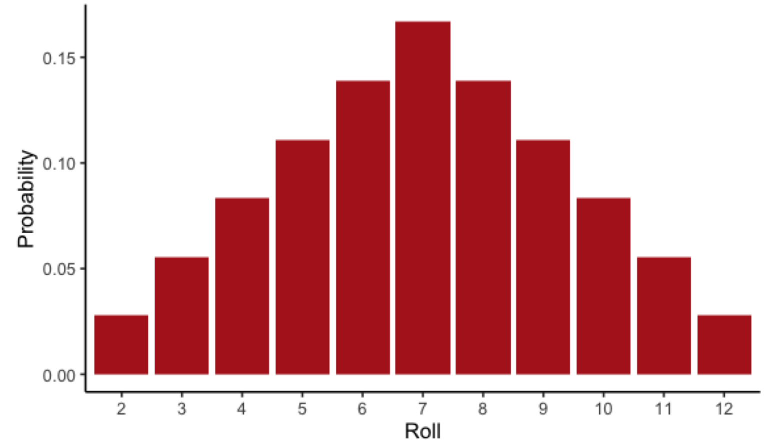
Probability distribution

A probability distribution describes the true relative frequency of all possible values of a random variable.

Probability distribution for the outcome of a roll of a die



Probability distribution for the sum of a roll of two dice



The addition principle

The addition principle: If two events A and B are mutually exclusive, then
Pr[A OR B] = Pr[A] + Pr[B]

The probability of a range

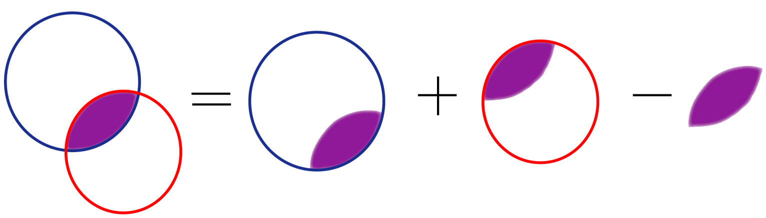
Pr[Number of green M&Ms ≥ 6] =
Pr[6 green] + Pr[7 green] + Pr[8 green]....

The probabilities of all possibilities add to 1.

Probability of *Not*

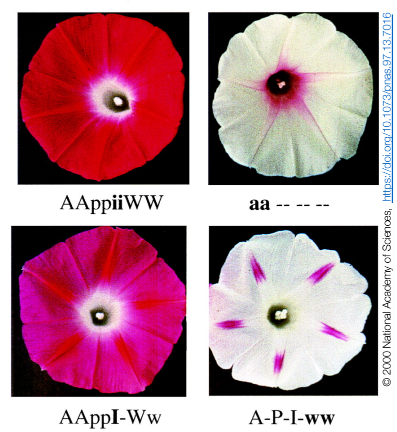
Pr[NOT *rolling a 2*] = 1 – Pr[*Rolling a 2*] = 5/6

General Addition Principle



General addition principle

Pr[*A* OR *B*] = Pr[*A*] + Pr[*B*] - Pr[*A* AND *B*].



e.g., Morning glory flowers can be white because of their *ww* genotype at the W locus and/or because of their *aa* genotype at the A locus

General addition principle

Pr[*A* OR *B*] = Pr[*A*] + Pr[*B*] - Pr[*A* AND *B*].

e.g., Morning glory flowers can be white because of their *ww* genotype at the W locus and/or because of their *aa* genotype at the A locus

If:

Pr[*ww*] = 0.1   Pr[*aa*] = 0.2

Pr *ww* AND *aa*] = 0.04

Pr[white flowers] = Pr[*ww*] + Pr[*aa*] – Pr[*ww* AND *aa]*

= 0.1 + 0.2 – 0.04

= 0.26

Multiplication principle

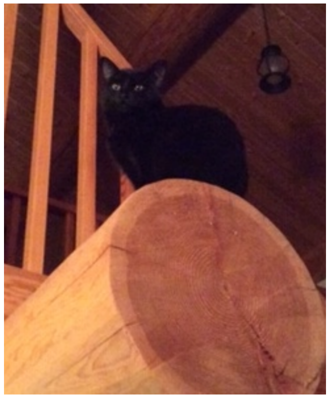
The multiplication principle: If two events A and B are independent, then:

Pr[A AND B] = Pr[A] × Pr[B]

Independence

Two events are independent if the occurrence of one gives no information about whether the second will occur.

Short-haired, black cats

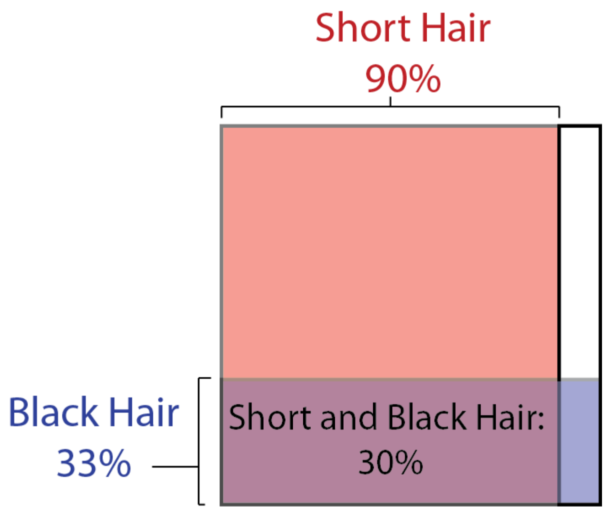


What is the probability that a cat is black- and short-haired?

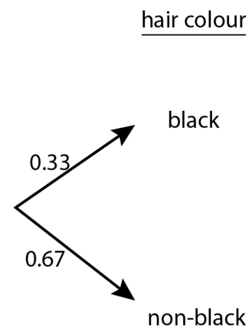
Pr[black hair] = 0.33   Pr[short-haired] = 0.90

Pr[black hair AND short hair] = 0.33 × 0.90

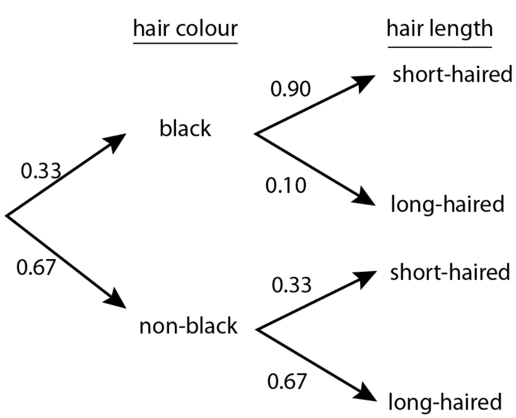
= 0.30.



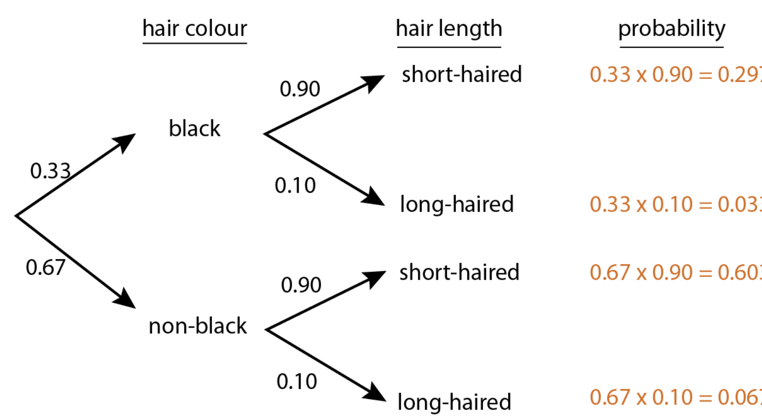
Probability trees



Phenotypes in two-child family



Phenotypes in two-child family



Short summary

The probability of A OR B involves addition.

Pr(A or B) = Pr(A) + Pr(B) if the two are mutually exclusive.

The probability of A AND B involves multiplication.

Pr(A and B) = Pr(A) Pr(B) if the two are independent

Dependent events

Variables are not always independent.

The probability of one event may depend on the outcome of another event.

Triple test and detection of trisomy 23

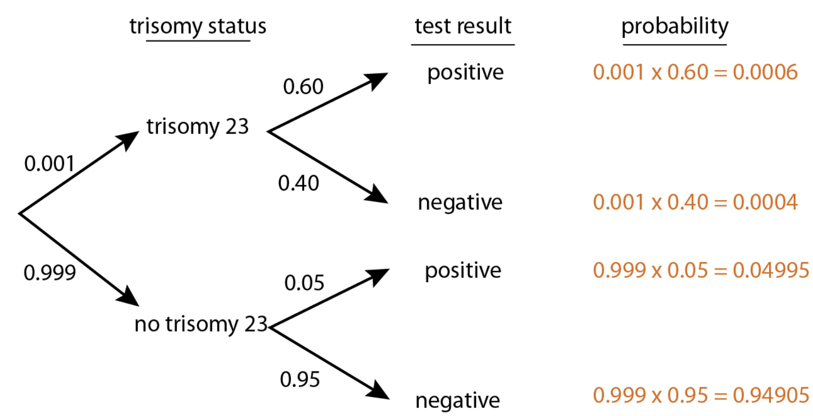
The triple test detects Down syndrome when it is present 60% of the time.

This means it has a 40% false negative rate of 40%.

The triple test gives a false positive 5% of the time (when the fetus does not have trisomy 23).

The probability that a fetus has trisomy 23 is 0.001.

Triple test outcomes



Are trisomy status and test results independent?

Pr(trisomy) = 0.001

Pr(positive test result) = 0.006 + 0.04995 = 0.05595

Pr(trisomy AND positive result) = 0.006 ≠
Pr(trisomy) × Pr(positive result) =
0.001 × 0.05595 = 0.000056

So these two events are NOT independent.

Conditional probability

The conditional probability of an event is the probability of that event occurring *given that* a condition is met.

Pr[X | Y]

Pr(*X* | *Y*) means the probability of *X* if *Y* is true.

It is read as "the probability of *X* given *Y*."

Pr(positive test | trisomy) = 0.60.

Law of total probability:

$$\Pr[X] = \sum_{All\ values\ of\ Y} \Pr[X|Y]\Pr[Y]$$

The probability of a positive test result is

Pr[positive result] =
Pr(positive result | trisomy) Pr(trisomy) +
Pr(positive result | no trisomy) Pr(no trisomy)

= 0.60 (0.001) + 0.05 (0.999) = 0.05055

The general multiplication rule

$$\Pr[A \text{ AND } B] = \Pr[A]\,\Pr[B \mid A].$$

The general multiplication rule

Pr[A AND B] = Pr[A] Pr[B | A].

Pr[A AND B] = Pr[B] Pr[A | B].

Therefore

Pr[A] Pr[B | A] = Pr[B] Pr[A | B].

Bayes' theorem

$$\Pr[A \mid B] = \frac{\Pr[B \mid A]\,\Pr[A]}{\Pr[B]}$$

Applying Bayes' theorem

For the triple test, what is the probability that a pregnancy with a positive result is affected by trisomy 23?

In other words:

Pr[trisomy 23 | positive result] = ?

$$\Pr[trisomy\ 23 \mid positive\ result] = \frac{\Pr[positive\ result \mid trisomy\ 23]\ \Pr[trisomy23]}{\Pr[positive\ result]}$$

$$\Pr[trisomy\ 23 \mid positive\ result] = \frac{\Pr[positive\ result \mid trisomy\ 23]\ \Pr[trisomy23]}{\Pr[positive\ result]}$$

We already know:

$$\Pr[positive\ result \mid trisomy\ 23] = 0.60$$

$$\Pr[trisomy23] = 0.0001$$

$$\Pr[trisomy\ 23 \mid positive\ result] = \frac{\Pr[positive\ result \mid trisomy\ 23]\ \Pr[trisomy23]}{\Pr[positive\ result]}$$

We need to know:

$$\Pr[positive\ result] = ?$$

$\Pr[trisomy\ 23 \mid positive\ result] =$

$$\frac{\Pr[positive\ result \mid trisomy\ 23]\ \Pr[trisomy23]}{\Pr[positive\ result]}$$

We need to know: (Law of total probability)

$\Pr[positive\ result] =$
$\Pr[positive\ result \mid trisomy\ 23]\ \Pr[trisomy23]$
$+ \Pr[positive\ result \mid no\ trisomy]\ \Pr[no\ trisomy]$
$= 0.60(0.001) + 0.05(0.999)$
$= 0.05595$

$\Pr[trisomy\ 23 \mid positive\ result] =$

$$\frac{\Pr[positive\ result \mid trisomy\ 23]\ \Pr[trisomy23]}{\Pr[positive\ result]}$$

$$= \frac{0.60 \times 0.001}{0.05595}$$

$= 0.011$