Combinatorics (counting tools)

The *mn* rule

If $a \in \{a_1,\ldots,a_m\}$ and $b \in \{b_1,\ldots,b_n\}$, then the number of different (ordered) pairs (a,b) is mn.

Example 7 Find the pr of getting at least one 6 on 2 rolls of a die.

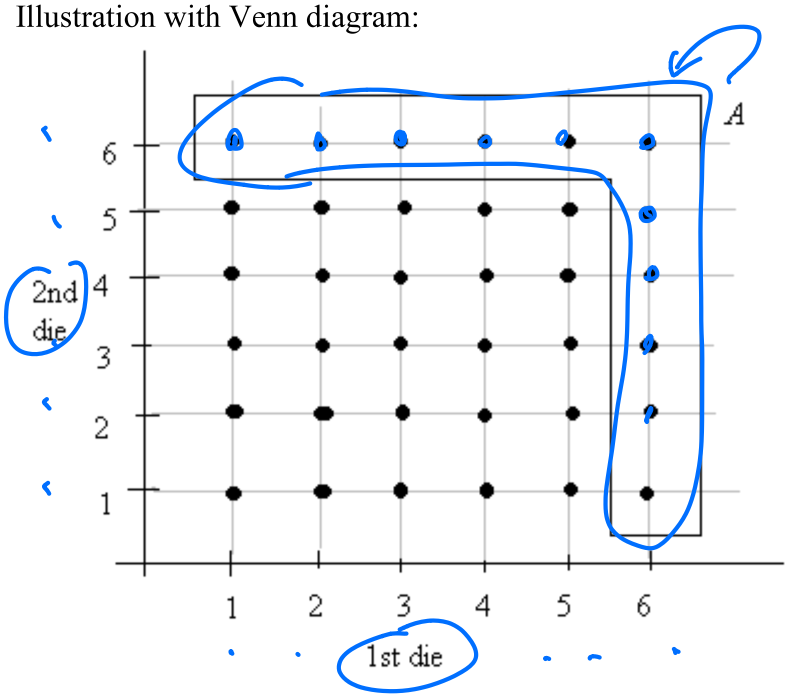
Let (5,3) denote a 5 coming up on the 1st roll and a 3 on the 2nd, etc, and abbreviate (a,b) by ab. Then the sample space is

$$S = \{11,12,13,\ldots,66\} = \{ab : a \in \{1,\ldots,6\}, b \in \{1,\ldots,6\}\}.$$

Hence (by the *mn* rule with $m = 6$ and $n = 6$), $n_S = 6 \times 6 = 36$.

Next let A = "At least one 6 comes up".
Then A ={16, 26, 36, 46, 56, 66, 65, 64, 63, 62, 61}, and so n_A =11.
Since all possible outcomes are equally likely, $P(A) = n_A / n_S$ = 11/36.

Illustration with Venn diagram:



11/36

Factorials

For n a positive integer, we define $n! = n(n-1)(n-2)\ldots(3)(2)(1)$.
We also define $0! = 1$.

Eg: $4! = 4(3)(2)(1) = 24$.

Theorem 5 The number of different possible arrangements of n distinct objects in a row is $n!$.

Example 8 How many different 3-letter words can be formed from the letters A, C, T (assuming that no letter can be used more than once).

$\boxed{3}\boxed{2}\boxed{1}$

There are $3! = 3(2)(1) = 6$ words (ACT, ATC, CAT, CTA, TAC, TCA).

Permutations

Defn

An *ordered* arrangement of objects is called a *permutation* (eg CAT above). The number of different permutations of r objects which can be formed from n distinct objects is denoted P_r^n.
(We say "n-permute-r". Other notations: $P(n,r)$, ${}^nP_r$, ${}_nP_r$)

Theorem 6 $P_r^n = \dfrac{n!}{(n-r)!}$ $\left(= \dfrac{n(n-1)...(n-r+1)\cancel{(n-r)!}}{\cancel{(n-r)!}} = n(n-1)...(n-r+1) \right)$.

Example 9 How many different words of length 3 can be formed from the letters A, C, T, E, W (assuming that no letter can be used more than once)?

There are $P_3^5 = \dfrac{5!}{(5-3)!} = \dfrac{120}{2} = 60$ words (ACT, ACE, ..., CAT, ..., TAC, ..., WET).

(To see this another way, there are 5 possibilities for the 1st letter in the word, then 4 possibilities for the 2nd, and finally 3 for the 3rd. Thus $P_3^5 = 5*4*3 = 60$.)

Permutations : ordered $\boxed{5 * 4 * 3} = \dfrac{5 * 4 * 3 * 2 * 1}{2 * 1} = \dfrac{5!}{(5-3)!}$

Combinations

An *unordered* collection of objects is called a *combination*.
The number of different combinations of r objects that can be selected from n distinct objects is denoted $\binom{n}{r}$.
(We say "n-choose-r". Other notations: $C(n,r)$, C_r^n, ${}^nC_r$, ${}_nC_r$.)

Theorem 7 $\dbinom{n}{r} = \dfrac{n!}{r!(n-r)!}$ $\left(= \dfrac{P_r^n}{r!} = \dfrac{n(n-1)...(n-r+1)}{r(r-1)...(r-r+1)} \right)$.

Example 10 How many different combinations of 3 letters can be formed from A, C, T, E, W (assuming that no letter can be used more than once)?

There are $\binom{5}{3} = \frac{5!}{3!2!} = \frac{120}{6(2)} = 10$ combinations

(ACT, ACE, ACW, ATE, ATW, AEW, CTE, CTW, CEW, TEW).

ACE, AEC, . . .

(Note: Each of these combinations corresponds to 3! = 6 permutations in Example 9. For instance, ACT corresponds to permutations ACT, ATC, CAT, CTA, TAC, TCA. Thus, another way to express the number of combinations is as $\binom{5}{3} = \frac{P_3^5}{3!} = \frac{60}{6} = 10$.)

A related question: How many different committees of size 3 can be selected from 5 people? The answer is the same, $\binom{5}{3}$ = 10. Think of the persons as labelled 1,2,3,4,5 & consider how many unordered sets of 3 numbers can be selected from these labels.

Note that another way to derive the same answer is to think about how many sets of 2 numbers can be selected from 1,...,5 (so as to form the 'non-committee'), ie $\binom{5}{2}$ = 10.

$\binom{5}{3} \quad \frac{5!}{3!2!} \quad \equiv \quad \frac{5!}{2!3!} \quad \binom{5}{2}$

Example 11 A committee of 5 is to be selected randomly from 12 people. What is the probability that it will contain the two oldest people?

The total number of committees is $\binom{12}{5} = \frac{12!}{5!7!} = 792$.

The number of committees with the 2 oldest people is $\binom{12-2}{5-2} = \binom{10}{3} = \frac{10!}{3!7!} = 120$.

(If the two oldest persons are definitely on the committee, then 3 more people need to be selected from the remaining 10.)

Hence the required probability is 120/792 = 5/33 = 0.152 (since all 792 committees are equally likely).

Example 12 Two separate committees of size 3 and 4 are to be selected from 15 people. How many different pairs of committees are possible?

There are $\binom{15}{3}$ possibilities for the 1st committee.

For each of these, there are $\binom{15-3}{4}=\binom{12}{4}$ possibilities for the 2nd committee.

So the number of possible pairs of committees is

$$\binom{15}{3}\binom{12}{4}=\frac{15!}{3!12!}\times\frac{12!}{4!8!}=\boxed{\frac{15!}{3!4!8!}}=225225.$$

Note 1: Alternatively, there are $\binom{15}{4}$ ways to choose the 2nd committee, and then $\binom{11}{3}$ ways to choose the 1st committee. So the number of possible pairs is

$$\binom{15}{4}\binom{11}{3}=\frac{15!}{4!11!}\times\frac{11!}{3!8!}=\boxed{\frac{15!}{3!4!8!}}=225225 \text{ (the same).}$$

Note 2: Yet again, there are $\binom{15}{8}$ ways to choose the 'non-committee', and then $\binom{7}{3}$ ways to select the 1st committee. So the number of possible pairs is

$$\binom{15}{8}\binom{7}{3}=\frac{15!}{8!7!}\times\frac{7!}{3!4!}=\boxed{\frac{15!}{3!4!8!}}=225225.$$

Note 3: (Theorem 2.3 in the text) The number of ways of partitioning n distinct objects into k distinct groups, containing $r_1,\dots,r_k$ objects, respectively is given by

$$\binom{n}{r_1 \;\dots\; r_k}=\binom{n}{r_1}\binom{n-r_1}{r_2}\binom{n-r_1-r_2}{r_3}\dots\binom{n-r_1-r_2-\dots-r_{k-2}}{r_{k-1}}=\boxed{\frac{n!}{r_1!\dots r_k!}},$$

where it is assumed that $r_1+\dots+r_k=n$.

In Example 12, $n = 15$, $r_1 = 3$, $r_2 = 4$ and $r_3 = 8$ (the number of persons *not* selected), so that the answer may be written $\binom{15}{3\;\;4\;\;8}=\binom{15}{3}\binom{12}{4}=225225$.

Note 4: If we allow for the possibility that $r_1+\dots+r_k<n$, we must generalise our definition of $\binom{n}{r_1 \;\dots\; r_k}$ as equal to $\binom{n}{r_1 \;\dots\; r_k \;\; r_{k+1}}=\frac{n!}{r_1!\dots r_k!r_{k+1}!}$, where $r_{k+1}=n-r_1-\dots-r_k$. We may then write $\binom{15}{3\;\;4\;\;8}$ as $\binom{15}{3\;\;4}$ or $\binom{15}{8\;\;3}$, etc.

Likewise, $\binom{n}{r}$ and $\binom{n}{n-r}$ are short for $\binom{n}{r\;\;n-r}$ and $\binom{n}{n-r\;\;r}$, respectively.

Example 6 (again) What's the pr of getting 2 H's on 5 tosses of a coin?

S = {HHHHH, HHHHT, ...}. How many sample points are there in S?
$n_S = 2^5 = 32$ by the *mn* rule (or, rather, an obvious extension of it).
Let A = "Get 2 H's". How many sample points are there in A?

Now A = {HHTTT, HTHTT, HTTHT, ..., TTTHH}.
12   1 3   1 4   ....   45 <---- 'position numbers'

We see that n_A is the number of combinations of 2 numbers that can be selected from 1,2,3,4,5, namely $\binom{5}{2} = 10$ (12, 13, 14, 15, 23, 24, 25, 34, 35, 45).

Thus n_A =10 and $P(A) = n_A / n_S$ =10/32 = 5/16 = 0.3125.

The bijection principle

In the solution to Example 6 above we made use of the *bijection principle*, which says that if there is a *bijection* (or *one-to-one-correspondence*) between two sets then those sets have the same number of objects. (The two sets were A = {HHTTT, HTHTT, HTTHT, ..., TTTHH} and {12, 13, ...,45}, with each set having 10 objects.)

To give another illustration of the bijection principle, consider the problem of counting the number of ways that three identical rings can be arranged on the five fingers of one hand, assuming that each finger can have any number of rings.

(+ ≡ 1 ≡ I)

To solve this problem we set up a bijection between arrangements of rings and arrangements of 0, 0, 0, +, +, +, + in a row. Here, each 0 stands for a ring and each + represents a division between two fingers. For example:

000++++	<----->	All 3 rings on Finger 1
00+++0+	<----->	Two rings on Finger 1 and one ring on Finger 4
++0+0+0	<----->	One ring on each of Fingers 3, 4 and 5, etc.

We see that the total number of arrangements of rings must be $\binom{7}{3} = 35$.

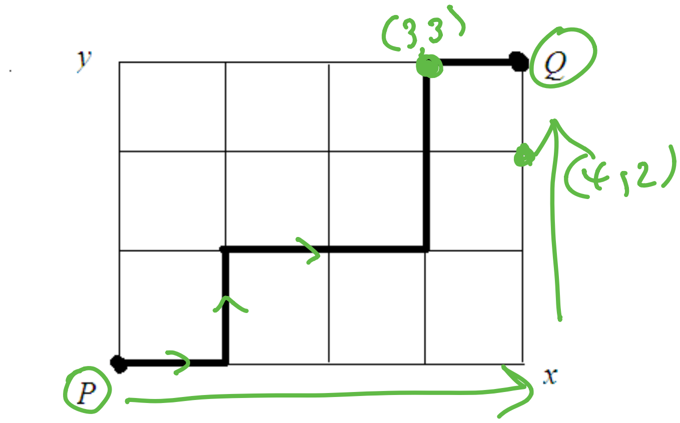
distinguishable

What if the rings are all different? Then each arrangement of rings can be represented by an arrangement of 1, 2, 3, +, +, +, + in a row. Now, if the numbers were 1, 2, 3, 4, 5, 6, 7, the answer would be 7!. But there are 4! ways that 4, 5, 6, 7 can be arranged in a row. So 7! is too large by a factor of 4! and the answer must be 7!/4! = 7*6*5 = 210.

This assumes ~~for~~ we are accounting for order on a finger eg (123)++++ ≠ (132)++++
what if the order does not matter? Then the Answer is $5^3 = 5\times5\times5 = 125$

Lattice paths

Consider the following *lattice* and all the possible *paths* from $P = (0,0)$ to $Q = (4,3)$, moving only *up* and *right* along the lines:



Each path must go up exactly 3 times and go right exactly 4 times, in any order. Some possibilities are RURRUUR (shown), RRRRUUU and RRRUURU.
We see that there is a one-to-one correspondence between paths and arrangements of 4 R's and 3 U's in a row. So the total number of paths must be $\binom{4+3}{3} = \binom{7}{3} = 35$.

Now consider the fact that each path must pass through either (3,3) or (4,2).

The number of paths going through (3,3) is $\binom{3+3}{3}$,

and the number going through (4,2) is $\binom{4+2}{2}$.

It follows that $\binom{4+3}{3} = \binom{3+3}{3} + \binom{4+2}{2}$. (Check: $\binom{6}{3} + \binom{6}{2} = 20 + 15 = 35$.)

Generalizing this logic, we now consider all paths from $P = (0,0)$ to $Q = (a,b)$ and get

$$\binom{a+b}{b} = \binom{(a-1)+b}{b} + \binom{a+(b-1)}{b-1}.$$

This can also be written $\boxed{\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1}}$, which is known as *Pascal's Identity*.

Note 1: Pascal's Identity can also be proved algebraically, as follows:

$$\binom{m}{k}+\binom{m}{k-1}=\frac{m!}{k!(m-k)!}+\frac{m!}{(k-1)!(m-k+1)!}$$

$$=\frac{m!(m-k+1)}{k!(m-k+1)!}+\frac{m!k}{k!(m-k+1)!}$$

$$=\frac{m!}{k!(m-k+1)!}((m-k+1)+k)$$

$$=\frac{(m+1)!}{k!((m+1)-k)!}=\binom{m+1}{k}.$$

3rd

Note 2: Another proof of Pascal's Identity is to consider the number of committees of k persons that can be formed from m men and 1 woman. The answer is obviously $\binom{m+1}{k}$. But this must be the same as the number of committees with k men (i.e., no women), which is $\binom{m}{k}$, plus the number with $k-1$ men, which is $\binom{m}{k-1}$.

Note 3: Pascal's Identity is a special case of *Vandermonde's Identity*:

$$\binom{m+w}{k}=\binom{m}{0}\binom{w}{k}+\binom{m}{1}\binom{w}{k-1}+\ldots+\binom{m}{k-1}\binom{w}{1}+\binom{m}{k}\binom{w}{0}.$$

This can be proved using lattice paths. But the simplest proof is to consider the number of committees of k persons that can be formed from m men and w women. The answer is obviously $\binom{m+w}{k}$. But the number of committees with no men is $\binom{m}{0}\binom{w}{k}$, the number with one man is $\binom{m}{1}\binom{w}{k-1}$, and so on. Summing up proves the result.

Note 4: In Vandermonde's Identity, k may be greater than m or w. For example,

?

$$\binom{3+1}{2}=\binom{3}{0}\binom{1}{2}+\binom{3}{1}\binom{1}{1}+\binom{3}{2}\binom{1}{0}=1\times 0+3\times 1+3\times 1=0+3+3=6.$$

In some cases the sample point method is impractical, even with assistance from combinatorics. This leads us to consider the second major strategy for computing probabilities.

The event composition method (2 steps)

1.	Express the event of interest A (say), as a composition (ie function) of other events (using unions, intersections, complementations, etc).
2.	Apply any relevant results to this composition in an attempt to compute $P(A)$.

Example 13 Two dice are rolled . Find the pr that the sum of the numbers which come up is at least 4.

1. Let A = "Sum is at least 4" and B = "Sum is less than 4".
We may write $A = \bar{B}$ (or $A = S - B$).
(A is expressed as a function of another event, B, or of other events, S and B).

2. Now $B = \{11, 12, 21\}$, so that $P(B) = n_B / n_S = 3/36 = 1/12$.
So $P(A) = \boxed{P(\bar{B}) = 1 - P(B)}$ (by Theorem 3 from earlier)
$= 1 - 1/12$
$= 11/12.$

We have here given an example involving one of the simplest forms of composition , namely complementation.

Another such example involving complementation is provided by an alternative solution to Problem 7 (Find the pr of getting at least one 6 on 2 rolls of a die).

Let A = "At least one six". Then $\bar{A}$ = "No sixes" = "Get 1 to 5 on each of the rolls", so that $n_{\bar{A}} = 5 \times 5 = 25$ (using the mn rule again) and hence $P(\bar{A}) = n_{\bar{A}} / n_S = 25/36$.

Thus $P(A) = 1 - P(\bar{A}) = 1 - 25/36 = 11/36$, as before.

This solution via the *event composition method* is only slightly simpler and easier than the solution via the *sample point method* used in Example 7 (where A was listed and all 11 of its elements counted). However, in many problems the event composition method will be very much simpler and easier to apply than the sample point method.

Let's now develop some more concepts and results that are useful for computing probabilities using the event composition method.

Conditional probability

For any two events A and B such that $P(B) > 0$,
the conditional probability of A given B is defined as

$$P(A \mid B) = \frac{P(AB)}{P(B)}.$$

If $P(B) = 0$, then $P(A|B)$ is undefined.

Example 14 A die is rolled. What is the pr that the number which comes up is even, given that it is greater than 3?

Let A = "An even number comes up" and B = "A number greater than 3 comes up".
Then $A = \{2,4,6\}$, $B = \{4,5,6\}$, $AB = \{4,6\}$, $P(AB) = 2/6$ and $P(B) = 3/6$ (> 0).
Hence $P(A|B) = P(AB)/P(B) = (2/6)/(3/6) = 2/3$.

Alternatively, $P(A|B) = n_{AB}/n_B = 2/3$.

Venn D.

Another interpretation: If someone rolled a die and we learn that a number greater than 3 came up, then we can be 66.7% confident that that number was even.

4 5 or 6 4 or 6

Yet another interpretation: Suppose we will roll a die repeatedly until a number greater than 3 comes up. Then the probability that the last number will be even is 2/3.

Yet another: We roll a die millions of times and each time write the number that comes up *if* it's more than 3. Then about 2/3 of the written numbers will be even.

iff

Independence

> Two events A and B are said to be *independent* if $P(AB) = P(A)P(B)$,
> in which case we write $A \perp B$.
> If $P(AB) \neq P(A)P(B)$, then A and B are *dependent* and we write $A \not\perp B$.

It can be shown that: If $P(A|B) = P(A)$ or $P(B|A) = P(B)$, then $A \perp B$.
If $P(A|B) \neq P(A)$ or $P(B|A) \neq P(B)$, then $A \not\perp B$.

Example 15 In Example 14, are A and B independent events?

$P(A)P(B) = (3/6)(3/6) = 9/36$ and $P(AB) = 2/6$. Thus $P(AB) \neq P(A)P(B)$.
So no, A and B are not independent; they are dependent.

{4, 6}

(Another solution: $P(A|B) = 2/3 \neq 3/6 = P(A) \Rightarrow A \not\perp B$.)

The multiplicative law of probabilitity (MLP)

> $P(AB) = P(A)P(B|A)$. $= P(B)P(A|B)$

(This follows directly from the definition of conditional pr.)

Example 16 Two cards are to be drawn from 5 white cards and 3 black cards.
Find the pr that a white card will be drawn first and then a black card.

1	2
1	2

1	2
2	1

Let A = "A white card is drawn first" and B = "A black card is drawn second".
Then $P(A) = 5/8$ and $P(B|A) = 3/7$.
Hence $P(AB) = P(A)P(B|A) = (5/8)(3/7) = 15/56 = 0.268$.

OR

$P(AB) = P(B)P(A|B) = \frac{3}{8} \times \frac{5}{7} = \frac{15}{56}$

The law of total probability (LTP)

> $P(A) = P(AB) + P(A\bar{B})$
> $= P(B)P(A|B) + P(\bar{B})P(A|\bar{B})$.

A useful corollary: $P(AB) = P(A) - P(A\bar{B})$. (*)

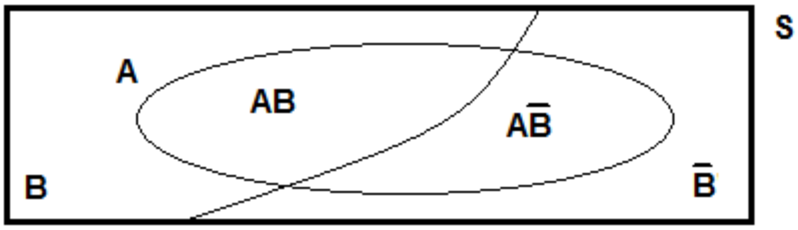
Proof of LTP: $P(A) = P(AS)$

$= P(A(B \cup \bar{B}))$

$= P((AB) \cup (A\bar{B}))$ by the distributive laws

$= P(AB) + P(A\bar{B})$ by Theorem 2 earlier (since $(AB)(A\bar{B}) = \varnothing$)

$= P(B)P(A \mid B) + P(\bar{B})P(A \mid \bar{B})$ by the MLP.



Example 17

We will consider a location where the following has been found to be true (approx.):

If it's cloudy tonight, the pr of rain tomorrow is 70%.
If it's *not* cloudy tonight, the pr of rain tomorrow is only 40%.
The pr of it being cloudy tonight is 20%.

Find the probability that it will rain tomorrow.

Let R = "Rain tomorrow" and C = "Clouds tonight".
Then $P(R|C) = 0.7$, $P(R \mid \bar{C}) = 0.4$, $P(C) = 0.2$ (and $P(\bar{C}) = 0.8$).

So: $P(CR) = P(C)P(R|C) = 0.2(0.7) = 0.14$
(the pr of clouds tonight and rain tomorrow)
$P(\bar{C}R) = P(\bar{C})P(R \mid \bar{C}) = 0.8(0.4) = 0.32$
(the pr of *no* clouds tonight and rain tomorrow).

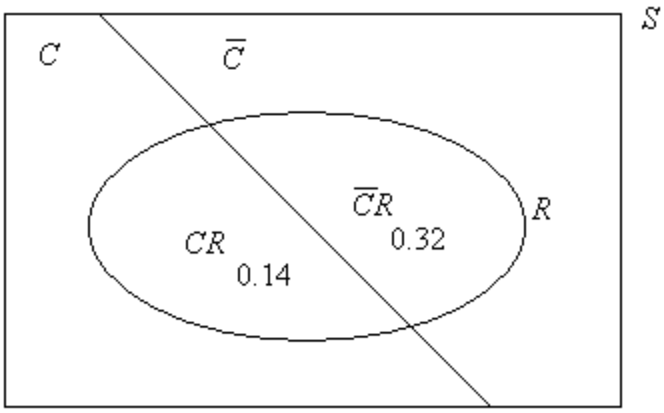
Thus the overall probability of rain tomorrow is
$P(R) = P(CR) + P(\bar{C}R) = 0.14 + 0.32 = 0.46.$

Equivalently, we could write: $P(R) = P(C)P(R \mid C) + P(\bar{C})P(R \mid \bar{C})$
$= 0.2(0.7) + 0.8(0.4) = 0.46.$

Note that the *unconditional* pr of rain tomorrow (46%) is an appropriately weighted average of the two *conditional* pr's (70% and 40%).

$P(R|C)$ $P(R|\bar{C})$

Illustration with Venn diagram:



The additive law of probability ALP

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

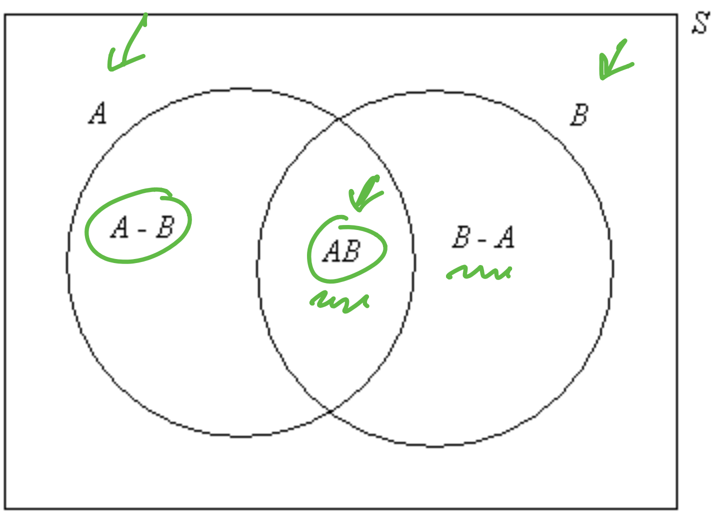
Proof: $A \cup B = (A - B) \cup (AB) \cup (B - A)$

(the union of three disjoint events, as shown in the Venn diagram below)

So $P(A \cup B) = P(A - B) + P(AB) + P(B - A)$

$= \{P(A - B) + P(AB)\} + \{P(B - A) + P(AB)\} - P(AB)$

$= P(A) + P(B) - P(AB).$



Example 18 Refer to Ex. 17. What's the pr it will be cloudy tonight or rain tomorrow? (or both)

$$P(C \cup R) = P(C) + P(R) - P(CR) = 0.2 + 0.46 - 0.14 = 0.52.$$

Note that this covers the possibility of clouds tonight *and* rain tomorrow.

So, what's the pr it will either be cloudy tonight or rain tomorrow, *but not both*?

$$P(C - R) + P(R - C) = P(C \cup R) - P(CR) = 0.52 - 0.14 = 0.38.$$

Bayes' rule (or theorem)

$\frac{P(BA)}{P(A)}$ by defn.

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

by MLP

$$= \frac{P(B)P(A \mid B)}{P(B)P(A \mid B) + P(\bar{B})P(A \mid \bar{B})}.$$

by LTP

(The two equalities here follow trivially from the multiplicative law of pr, the LTP and the defn of conditional pr.)

before A after A

We call $P(B)$ the *prior probability* of B, and $P(B|A)$ the *posterior probability* of B. Bayes' rule shows that the posterior is the prior multiplied by a factor $P(A \mid B) / P(A)$.

Example 19 Refer to Ex. 17.
Suppose tomorrow has come and it's raining.
What's the pr that it was cloudy last night?

$$P(C \mid R) = \frac{P(C)P(R \mid C)}{P(R)} = \frac{0.2(0.7)}{0.46} = 0.304$$

(the posterior pr that it was cloudy last night).

$P(C|R) > P(C)$
$.304 > .2$
$(\checkmark)$

(Alternatively, $P(C \mid R) = \frac{P(CR)}{P(R)} = \frac{0.14}{0.46} = 0.304$.)

Observe that the posterior pr is *higher* than the prior pr, $P(C) = 0.2$.
This makes sense because being cloudy is associated with a *higher* chance of rain.

Now suppose that tomorrow has come and it's *not* raining. What then is the pr of it being cloudy last night? We suspect that this pr is *lower* than 0.2. Let's find out.

$$P(C\mid \bar{R}) = \frac{P(C\bar{R})}{P(\bar{R})} = \frac{P(C)P(\bar{R}\mid C)}{P(\bar{R})} = \frac{P(C)\{1-P(R\mid C)\}}{1-P(R)} = \frac{0.2(1-0.7)}{1-0.46} = 0.111, \quad < .2$$

(✓)

which is indeed lower than $P(C) = 0.2$.

Summary of probabilities in Example 17 to 19:

Event	*Probability*	*Description*
C	0.2	(prior pr. of clouds tonight)
C'	0.8	(prior pr. of no clouds)
R|C	0.7	(conditional pr. of rain, given clouds)
R|C'	0.4	(conditional pr. of rain, given no clouds)
R	0.46	(unconditional pr. of rain tomorrow)
CR	0.14	
C|R	0.14/0.46 = 0.304	(posterior pr. of clouds, given rain)
CR'	0.06	
C|R'	0.06/0.54 = 0.111	(posterior pr. of clouds, given no rain)
C'R'	0.48	
etc.		

Venn diagram: