2) Outcomes, events and probability

2.1 Sample Spaces

Def 2.1 An <u>experiment</u> is <u>anything with</u> a <u>set of possible outcomes</u>

The set of all possible outcomes is the <u>sample space</u> of the experiment.

Notation: Ω represents the sample space.

Example 2.2: Tossing a coin, <u>once</u>
Outcomes are: Tails : T
Heads : H

Sample space $\Omega = \{H, T\}$

Tossing a coin <u>twice</u>:

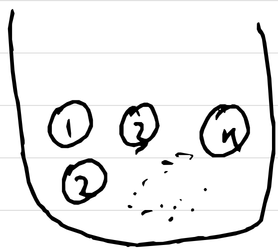
Sample space $\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$

↑
ordered pairs
↓
order is important

(H, T) means heads first, tails second.

Example 2.3 Experiment consists of drawing balls from a bag containing balls numbered:

$1 \ldots n$.



Then

$$\Omega = \{\text{all permutations of set } (1, \ldots n)\}$$

For example when $n=3$

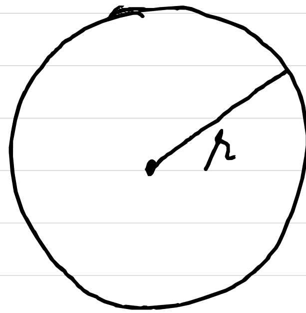
$$\Omega = \{ (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1) \}$$

Notation: We use $|E|$ to denote size of set (Also known as cardinality).

In Example 2.2, the second sample space, $|\Omega| = 4$

In example 2.3, $|\Omega| = n!$

Example 2.4 Throwing a dart at a board of radius r.



Measuring dart's distance d from the centre, the sample space consists of all distances between 0 and r. i.e.

$$\Omega = [0, r] = \{d \mid 0 \leq d \leq r\}$$

$|\Omega| > |\mathbb{N}| \Rightarrow$ sample space Ω is un-countable and infinite

2.2 Events

An event is a set containing a number of possible outcomes.

The event is said to have occured if the realised outcome is among the outcomes contained in the event.

Example 2.5: Let Ω be sample space for tossing coin twice.

$$\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$$

The event that exactly one heads is present is

$$\{(H,T), (T,H)\}$$

The event that the second toss is heads is

$$\{(H,H), (T,H)\}$$

Note: Events are always sets.

An event E is always subset of sample space

$$E \subseteq \Omega$$

The entire sample space can also be a subset hence an event.

$$\Omega \subseteq \Omega$$

Ω is called the sure event because it contains all possible outcomes.

The empty set ϕ is a subset of any set, hence

$$\phi \subseteq \Omega .$$

ϕ is called the impossible event.

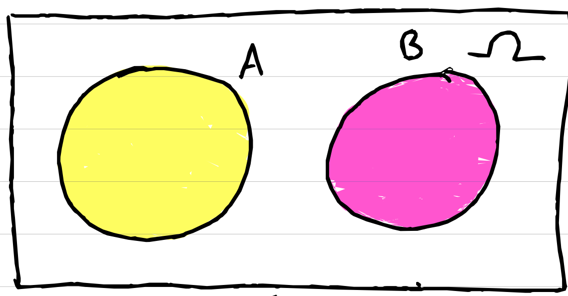
We can create new more complicated events from simple ones using set operations.

- $A \cup B$ (A "or" B): to be the event consisting of all outcomes belonging to either A or B or both A and B.

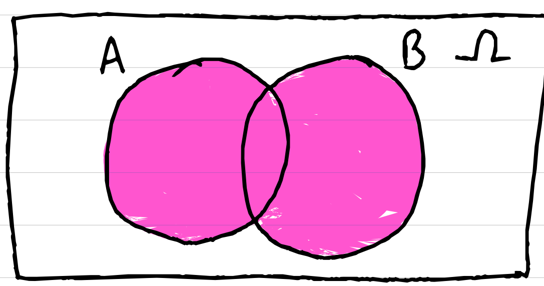
- $A \cap B$ (A "and" B): to be the event consisting of all outcomes that belong to both A and B.
  (Intersection of A and B)

• A^c ("not" A): to be the complement of A, that is, A^c contains all outcomes which are not members of A.

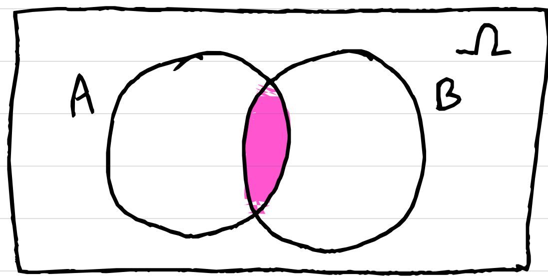
Using Venn diagrams to visualise set operations:



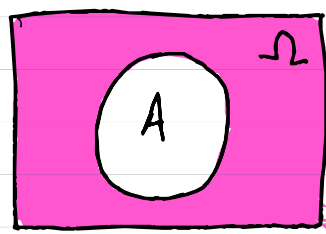
Just A and B



$A \cup B$



$A \cap B$



Example 2.6: In the setting of example 2.3,
Let $n = 3$,

A = number two is drawn out of bag first

B = number 3 is drawn second

$A = \{(2,1,3), (2,3,1)\}$

$B = \{(1,3,2), (2,3,1)\}$

$A \cup B = \{(1,3,2), (2,1,3), (2,3,1)\}$

$A \cap B = \{(2,3,1)\}$

$A^c = \{(1,2,3), (1,3,2), (3,1,2), (3,2,1)\}$

Note: $\cdot\ \Omega^c = \phi \quad \cdot\ \phi^c = \Omega \quad \cdot\ (E^c)^c = E$

Also for any events A and B,

$$\phi \subseteq A \cap B \subseteq A \subseteq A \cup B \subseteq \Omega$$

and so

$$0 = \phi \leq |A \cap B| \leq |A| \leq |A \cup B| \leq |\Omega|$$

Also define unions and intersections in the same way:

for example, for E_1, E_2, E_3,

$$E_1 \cup E_2 \cup E_3$$

represent atleast one of E_1, E_2, E_3 occur.

A convenient way to write expressions involving an arbitrary number of events is to introduce an index set I and write collection of events as

$$\{E_i \mid i \in I\}$$

For example: for three events E_1, E_2, E_3, could be written as

$$\{E_i \mid i \in I\}$$

with $I = \{1, 2, 3\}$ and write

$$E_1 \cup E_2 \cup E_3 \text{ as } \bigcup_{i \in I} E_i$$

Thus

- $\bigcup_{i \in I} E_i$ is the event that consists of all outcomes that belong to atleast one of the events in the collection.

- $\bigcap_{i \in I} E_i$ is the event that consists of all outcomes that belong to all of the events in the collection.

An <u>infinite</u> <u>countable</u> collection

$$E_1, E_2, \ldots$$

could be written as $\{E_i \mid i \in \mathbb{N}\}$. In that case there is an alternative notation:

$$\bigcup_{i \in \mathbb{N}} E_i = \bigcup_{i=1}^{\infty} E_i$$

<u>Proposition 2.7</u> (De Morgan's law):
For any countable collection of events, $\{E_i \mid i \in I\}$

$$\left(\bigcup_{i \in I} E_i\right)^c = \bigcap_{i \in I} E_i^c$$

$$\left(\bigcap_{i \in I} E_i\right)^c = \bigcup_{i \in I} E_i^c$$

proof look at lecture notes for proof.

By setting $I = \{1, 2\}$ and $E_1 = A$ and $E_2 = B$, we obtain:

Corollary 2.8 Let A and B be events. Then

$$\boxed{\begin{aligned}(A \cup B)^c &= A^c \cap B^c \\ (A \cap B)^c &= A^c \cup B^c\end{aligned}}$$

Proposition 2.9 The following rules hold for set operations for all events,

Let E, F, G be all events.

Commutativity:

$$\boxed{\begin{aligned}E \cup F &= F \cup E \\ E \cap F &= F \cap E\end{aligned}}$$

Associativity:

$$\boxed{\begin{aligned}(E \cup F) \cup G &= E \cup (F \cup G) \\ (E \cap F) \cap G &= E \cap (F \cap G)\end{aligned}}$$

Distributivity:

$$\boxed{\begin{aligned} E\cap(F\cup G) &= (E\cap F)\cup(E\cap G) \\ E\cup(F\cap G) &= (E\cup F)\cup(E\cup G) \end{aligned}}$$

proof: See Math Skills 1.

2.2.1 Event Space and σ algebras

A σ-algebra is a set that is closed under union, intersection and complement.

Defn 2.10 A σ-algebra $\mathcal{F}$ on a sample space Ω is a set of subsets of Ω such that:

1) $\Omega \in \mathcal{F}$ and $\phi \in \mathcal{F}$

2) If $E \in \mathcal{F}$ then $E^c \in \mathcal{F}$

3) If $\{E_i \mid i \in I\}$ is a countable subset of $\mathcal{F}$ then

$$\bigcup_{i\in I} E_i \in \mathcal{F} \quad \text{and} \quad \bigcap_{i\in I} E_i \in \mathcal{F}$$

The set of events for an experiment is called <u>σ-algebra</u>.

σ-algebra is also called <u>event space</u>.

σ-algebra summarises the following:

- Ω is the <u>certain/sure</u> event.
- ϕ is the impossible event.
- E is an event $\Rightarrow E^c$ is also an event.
- countable intersections and unions of events are also events.

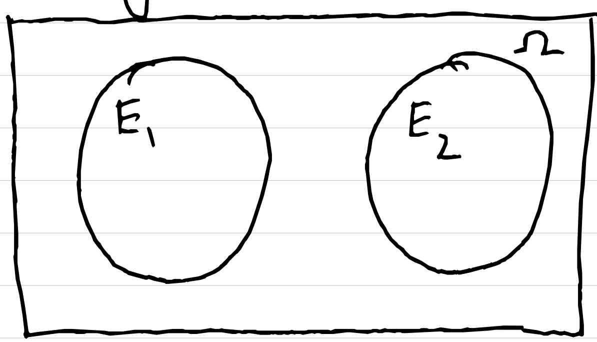
The <u>power set $P(\Omega)$ is</u> the <u>σ-algebra</u> <u>for Ω</u>

(However when Ω is uncountably infinite, the power set is too large and have to work with smaller σ algebra).

Defn 2.11 A collection of $\{E_i \mid i \in I\}$ of events is called <u>disjoint</u> or <u>mutually exclusive</u> if no two events share a common element. i.e.

$$E_i \cap E_j = \emptyset \quad \forall\, i \neq j \in I.$$

Venn diagram for E_1, E_2 (two events).



$$E_1 \cap E_2 = \emptyset$$

For example $\{1,2\}$, $\{3,4\}$ are disjoint.

The sets $\{1,2\}$, $\{3,4\}$, $\{4,5\}$ are <u>not</u> disjoint as last two sets share element 4.

2.3 Probability

Defn 2.12 (Axioms of probability):
Let Ω be the sample space of an experiment.
Let $\mathcal{F}$ be the σ-algebra of events.
A probability function P assigns to each event $E \in \mathcal{F}$ a real number $P(E)$ s.t.

(P1) $P(E) \in [0,1]$

(P2) $P(\Omega) = 1$

(P3) If $\{E_i \mid i \in I\}$ is a countable disjoint collection of events then

$$P\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} P(E_i)$$

The number $P(E)$ is called the probability that E occurs

The triple $(\Omega, \mathcal{F}, P)$ is a probability space

Note: (P3) contains a special case:

$$P(A \cup B) = P(A) + P(B) \text{ if } A \cap B \neq \phi.$$

Theorem 2.13: Let $(\Omega, \mathcal{F}, P)$ be the probability space. Then for any events $A, B \in \mathcal{F}$

(P4) $P(A^c) = 1 - P(A)$

(P5) $P(\phi) = \phi$

(P6) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(P7) If $A \subseteq B$ then $P(A) \leq P(B)$

proof: (P4):

Observe that $A \cap A^c = \phi$ and $A \cup A^c = \Omega$

$$P(\Omega) = 1 \qquad \text{by (P1)}$$

$$\Rightarrow P(A \cup A^c) = 1$$

$$\Rightarrow P(A) + P(A^c) = 1 \qquad \text{by (P3)}$$

$$\Rightarrow P(A^c) = 1 - P(A). \text{ Thus}$$

$$\boxed{P(A^c) = 1 - P(A)}$$

(P5):

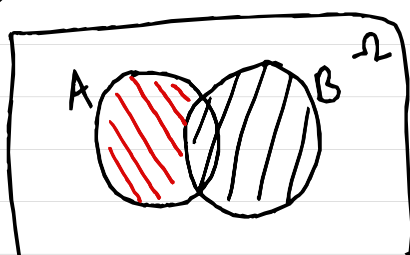
Observe that $\emptyset = \Omega^c$.

$$
\begin{aligned}
P(\emptyset) &= P(\Omega^c) \\
&= 1 - P(\Omega) && \text{by (P4)} \\
&= 1 - 1 && \text{by (P1)} \\
&= 0
\end{aligned}
$$

Thus

$$\boxed{P(\emptyset) = 0}$$

(P6):



$A \cup B = B \cup (A \cap B^c)$ and $B \cap (A \cap B^c) = \emptyset$.
So

$$P(A \cup B) = P(B) + P(A \cap B^c) \qquad \text{by (P3)}$$

$\hookrightarrow (\ast 1)$

$A = (A \cap B) \cup (A \cap B^c)$ and $(A \cap B) \cap (A \cap B^c) = \phi$.
So

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

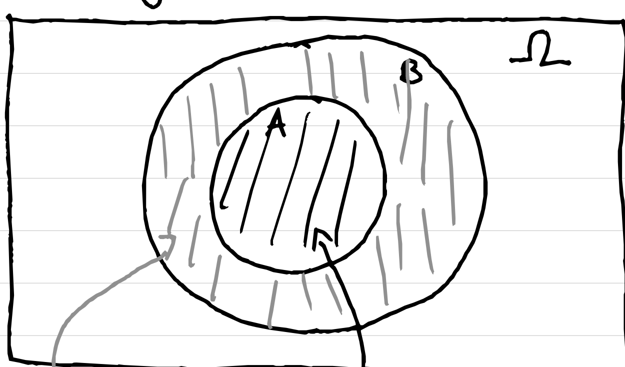
$\Rightarrow$

$$P(A \cap B^c) = P(A) - P(A \cap B). \qquad (*2)$$

Substituting $(*2)$ into $(*1)$ we get

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

(P7):
$A \subseteq B$
Venn diagram



$B = (B \cap A^c) \cup A$ and $A \cap (B \cap A^c) = \phi$.

So

$$P(B) = P((B \cap A^c) \cup A)$$

$$= P(B \cap A^c) + P(A) \qquad \text{by } (P3)$$

Since $P(B \cap A^c) \geq 0$ by (P1),

$$\boxed{P(B) \geq P(A)}$$

$\blacksquare$

__Example 2.14__ The probability space of a football match is $(\Omega, \mathcal{F}, P)$.

$\Omega = \{\text{Win}, \text{Draw}, \text{Lose}\} = \{W, D, L\}$

As σ-algebra $\mathcal{F}$ of events is the __power set__ of Ω,

$$\mathcal{F} = \mathcal{P}(\Omega) = \{\emptyset, \{W\}, \{D\}, \{L\}, \{W, D\}, \{W, L\}, \{D, L\}, \Omega\}$$

A probability function P assigns to each event $\mathcal{F}$ a real number.

A possible assignment is as follows:

Event	Probability
$\emptyset$	0
$\{W\}$	0.6
$\{D\}$	0.1
$\{L\}$	0.3
$\{W, D\}$	0.7
$\{W, L\}$	0.9
$\{D, L\}$	0.4
Ω	1

Example 2.15: Revisitting 2.4 experiment:
Sample space is

$$\Omega = [0, 1]$$

where we have chosen $r = 1$.
Next we need to chose the event space.
We cannot chose the power set, as it is uncountable.
The event space would also need to contain all subsets of $[0,1]$ that can be obtained by taking countable unions, intersections and complements, otherwise it would not be a σ-algebra.

The <u>smallest σ-algebra</u> containing all intervals is known as the <u>Borel algebra</u>.

So we <u>choose</u> $\underline{\mathcal{F}}$ to be the <u>Borel Algebra</u> on $[0,1]$

A probability function is now uniquely determined by specifying the probability of closed intervals

For example choose

$$P([a,b]) = b-a.$$

This leads to the probability function known as <u>Borel measure</u>.

An interesting bit of this example is that the probability that the dart lands at a particular distance d is zero for any d:

$$P(\{d\}) = P([d,d])$$

$$= d-d$$

$$= 0.$$

This does not mean that it is impossible for the dart to land at a certain distance d. In fact it must land at a certain distance, eventhough the probability of landing at any distance is zero.

The important lesson is that an event that has probability zero does not have to mean an impossible event. i.e.

$$P = 0 \not\Rightarrow \text{impossible} \qquad \text{but}$$

$$\text{impossible} \Rightarrow P = 0$$

Similarly

$$P = 1 \not\Rightarrow \text{certain} \qquad \text{but}$$

$$\text{certain} \Rightarrow P = 1$$

$P(A \cap B) = 0 \not\Rightarrow$ A and B are disjoint or mutually exclusive.

A and B are disjoint or mutually exclusive. $\Rightarrow P(A \cap B) = 0$

Theorem 2.16: Suppose the sample space Ω contains exactly n outcomes, $|\Omega| = n$
Let the event space $\mathcal{F}$ contain all subsets of Ω.
Set

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

for any event $E \in \mathcal{F}$.
Then P is a probability function.

proof: We need to show P satisfies the probability axioms.

(P1):
The number of elements of any event E must be between 0 and n:

$$0 \le |E| \le n,$$

it follows that

$$0 \le \frac{|E|}{n} = P(E) \le 1$$

So

$$P(E) \in [0,1]$$

satisfying (P1)

$(P2)$:

Since $|\Omega| = n$, if $E = \Omega$,

$$P(\Omega) = \frac{|\Omega|}{n} = \frac{n}{n} = 1$$

$\Rightarrow$

$$P(\Omega) = 1$$

satisfying $(P2)$

$(P3)$:

For disjoint sets $E_1, E_2, E_3, \dots, E_k$,

$|E_1 \cup E_2 \cup E_3 \cup \dots \cup E_k| = |E_1| + |E_2| + \dots + |E_k|$

Hence

$$\begin{aligned} P(E_1 \cup E_2 \cup \dots \cup E_k) &= \frac{|E_1 \cup \dots E_k|}{n} \\ &= \frac{|E_1| + |E_2| + \dots + |E_k|}{n} \\ &= \frac{|E_1|}{n} + \frac{|E_2|}{n} + \dots + \frac{|E_k|}{n} \\ &= P(E_1) + P(E_2) + \dots + P(E_k) \end{aligned}$$

satisfying $(P3)$. $\blacksquare$

Experiments where outcomes are not equally likely are not covered by Theorem 2.16

Example 2.17: Experiment consisting of drawing one ball at random.

Bag contains:
- 4 red balls
- 6 green balls
- 3 blue balls.

The set of possible outcomes are

$$\Omega = \{red, green, blue\} = \{r, g, b\}$$

The probabilities of these outcomes are

$$p_r = P(\{red\}) = 4/13$$

$$p_g = P(\{green\}) = 6/13$$

$$p_b = P(\{blue\}) = 3/13$$

The probability for all events can be calculated as

Event	P(Event)
$\emptyset$	0
{red}	$p_r = 4/13$
{green}	$p_g = 6/13$
{blue}	$p_b = 3/13$
{red, green}	$p_r + p_g = 10/13$
{red, blue}	$p_r + p_b = 7/13$
{green, blue}	$p_g + p_b = 9/13$
{red, green, blue} $= \Omega$	$p_r + p_g + p_b = 13/13 = 1$

In case of a countable sample space, once probabilities of <u>elementary events</u> (events containing single outcomes have been assigned, all other probabilities can be deduced from those. Only restriction is that <u>all elementary probabilities must add upto 1.</u>

This is formulated as a theorem on next page.

Theorem 2.18: Let Ω be a countable sample space. Let event space $\mathcal{F}$ be the power set of Ω:

$$\mathcal{F} = \mathcal{P}(\Omega)$$

Choose a set $\{p_w \mid w \in \Omega\}$ of nonnegative real numbers satisfying $\sum_{w \in \Omega} p_w = 1$

For any event $E \in \mathcal{F}$ define

$$P(E) = \sum_{w \in E} p_w$$

Then P is a probability function on $\mathcal{F}$ and thus $(\Omega, \mathcal{F}, P)$ is a probability space.

proof: We need to check conditions (P1), (P2), (P3).

(P1): Because all probabalities are sums of elementary probabalities p_w and

$$\sum_{w \in \Omega} p_w = 1,$$

Any probability is nonnegative and no larger than 1.

(P2): We have

$$P(\Omega) = \sum_{\omega \in \Omega} p_\omega = 1$$

(P3): This uses the fact that the sum over a disjoint union of index of sets can be split up into a sum over individual sums:

$$P\left(\bigcup_{i \in I} E_i\right) = \sum_{\omega \in \bigcup_{i \in I} E_i} p_\omega$$

$$= \sum_{i \in I} \sum_{\omega \in E_i} p_\omega$$

$$= \sum_{i \in I} P(E_i)$$

$\blacksquare$

The special case where all elementary probabilities are equal, Theorem 2.18 becomes Theorem 2.16.

2.4 Product of Sample spaces:

Suppose we perform two experiments with probability spaces $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$.

Then the possible outcomes of the combined experiment can be described by ordered pairs of the outcomes of the individual experiments and thus the combined sample space is

$$\boxed{\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) \mid \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}}$$

We have seen this in form in Example 2.2: Tossing a coin twice. This can be seen as a combination of two experiments each consisting of one flip of coin.
Individual sample spaces are

$$\Omega_1 = \{H, T\} \quad \text{and} \quad \Omega_2 = \{H, T\}$$

$$\Omega = \Omega_1 \times \Omega_2 = \{(H,H), (H,T), (T,H), (T,T)\}$$

Note: $|\Omega| = |\Omega_1| \cdot |\Omega_2|$.

If Ω_1 and Ω_2 is countable, so is Ω.
$\Rightarrow$
we can choose $\mathcal{F}$ as the powerset of Ω.
We can define the probability function of $\mathcal{F}$ as

$$\boxed{P(\{(\omega_1,\omega_2)\}) = P_1(\{\omega_1\}) \cdot P_2(\{\omega_2\}) \ \forall (\omega_1,\omega_2) \in \Omega.}$$

The probabilities of composite events are then determined as in Theorem 2.18. The thm tells us that the $(\Omega, \mathcal{F}, P)$ we have constructed in this way is a valid probability space, provided the sum of all elementary events' probability is 1.
Lets check this is indeed the case: (next page)

$$\sum_{(\omega_1,\omega_2)\in\Omega} P(\{(\omega_1,\omega_2)\}) = \sum_{\omega_1\in\Omega}\sum_{\omega_2\in\Omega} P(\{\omega_1,\omega_2\})$$

$$= \sum_{\omega_1\in\Omega_1}\sum_{\omega_2\in\Omega_2} P_1(\{\omega_1\})\cdot P_2(\{\omega_2\})$$

$$= \sum_{\omega_1\in\Omega_1} P(\{\omega_1\})\cdot \sum_{\omega_2\in\Omega_2} P(\{\omega_2\})$$

$$= P_1(\Omega_1)\cdot P_2(\Omega_2) = 1.1 = 1$$

summing over all pairs of outcomes is the same as summing over all possible pair of outcomes of one experiment and then summing over all outcomes of the other experiment.

The probability function of the joint sample space is the appropriate choice when the realised outcome of one experiment has no influence on the other experiment.

↳ important not to use product in other cases.

→ experiments are independant

The construction of probability space of two experiments can be generalised to n experiments:

$$\boxed{\Omega = \Omega_1 \times \dots \times \Omega_n = \{(\omega_1, \dots \omega_n) \mid \omega_i \in \Omega_i, i = 1 \dots n\}}$$

When all experiments are independant of each other:

$$P(\{\omega_1, \dots \omega_n\}) = P_1(\{\omega_1\}) \dots P_n(\{\omega_n\})$$

Example 2.19 (Chevalier de Méré):
Consider a game consisting of throwing a pair of dice 24 times in which you win if there is atleast one double 6 among the 24 throws, and you lose otherwise.
what is the probability of winning.

Solution: Sample space for a single throw of a pair of dice:

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

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

$$= 36$$

$\Rightarrow$

$$|\Omega| = 36$$

The probability of rolling a double 6 in one round:

$$P(\{6,6\}) = 1/36$$

The probability of not rolling a double 6 is therefore

$$P = (\{6,6\}^c) = 1 - 1/36$$

$$= 35/36$$

Individual throws are independant.
The probability of losing the game, i.e. in each of the 24 rounds not rolling a double 6 is the product of the probability of not rolling double 6 in first round times the probability of not rolling a double 6 in second round ...

$$P(L) = P_1(\{6,6\}^c) . P_2(\{6,6\}^c) . \ldots P_{24}(\{6,6\}^c)$$

$$= \left(\frac{35}{36}\right)^{24}$$

The probability of winning then is

$$P(W) = 1 - P(L)$$

$$= 1 - \left(\frac{35}{36}\right)^{24}$$

$$\approx 0.49$$

Appendix:

- Trick:

Any event D can be split into 2 disjoint components by choosing an arbitrary other event A and writing:

$$D = \Omega \cap D = (A \cup A^c) \cap D$$

$$= (A \cap D) \cup (A^c \cap D)$$

$$\Rightarrow \quad D = (A \cap D) \cup (A^c \cap D)$$