Probability

Definition:

- **Random** phenomenon: cannot be predicted with certainty in advance.
- **Outcome:** single observed result of random phenomenon.
- **Sample space:** set of all possible outcomes.
- **Empty set:** set containing no outcomes.
- **Event:** subset of sample space.
- The **probability** of an event is a number $0 \le p \le 1$ that describes how likely it is that the event occurs.
  - An event of probability 1 will happen for sure;
  - An event of probability 0 will certainly not happen.

  $P(S) = 1$, as the sample space includes all possibilities.
  $P(\emptyset) = 0$, as $\emptyset$ contains no possibilities.

- Some probabilities can be calculated, others can be found experimentally as long run proportions.
- They can be added, provided they are disjoint (mutually exclusive).

Example:

The probability that a random student obtains grades in MATH223 are:

F	P	C	D	HD
0.2	0.35	0.2	0.15	0.1

Let E denote the event $\{C, D, HD\}$ ("credit or better").

$P(E) = 0.2 + 0.15 + 0.1 = 0.45$

This is valid because events $\{C\}$, $\{D\}$, and $\{HD\}$ are disjoint (non-overlapping).

- If all outcomes are equally likely, then:

  $P(A) = \frac{|A|}{|S|}$, where $|X|$ is the number of outcomes in set X

Example:

A coin is tossed twice. The sequence of heads and tails is record. $S = \{HH, HT, TH, TT\}$. Let $E = \{HH, TT\}$ denote the event "same result for both tosses." Since all 4 outcomes have equal probability,

$$P(E) = \frac{|E|}{|S|} = \frac{2}{4} = \frac{1}{2}$$

Exercise:

2 fair dice are rolled. What is the probability that the sum of faces is 4?

A:

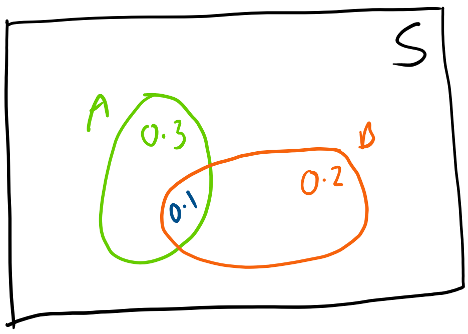
	1	2	3	5	6	TOTAL
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{6}$
TOTAL	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

All 36 events have equal probability, and 3 of them meet our needs.
Let $B = \{(1,3), (2,2), (3,1)\}$. Then,

$$P(B) = \frac{|B|}{|S|} = \frac{3}{36} = \frac{1}{12}$$

Venn Diagrams

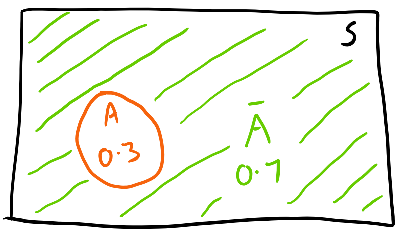
- A Venn Diagram represents the sample space and all events.
- Probabilities are represented as areas.



Complement

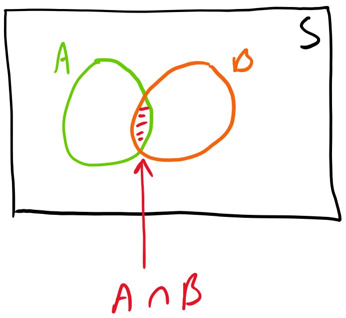
- The complement of A, denoted by A^c or $\overline{A}$, is the set of all outcomes not in A.

$$P(\overline{A}) = 1 - P(A)$$



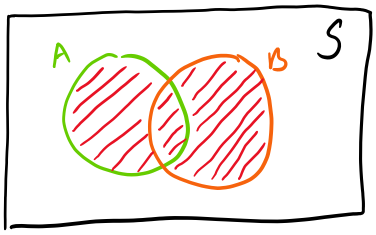
Intersection

- The intersection $A \cap B$ is the event that both A and B occur.



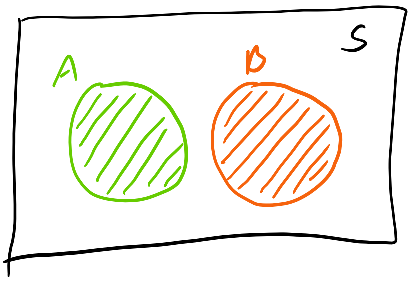
Union

- The union $A \cup B$ is the event that A or B (or both) occurs.



Disjoint

- Two events A and B are **disjoint** (cannot occur simultaneously) if $A \cap B = \emptyset$.



- For disjoint events A and B

$$P(A \cup B) = P(A) + P(B)$$
$$P(A \cap B) = 0$$

Conditional Probability

- The conditional probability of event A given that B has occurred is denoted by $P(A|B)$.
- In general (disjoint or not), $P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$.
- That is, for A and B both to happen, one event happens, and then given that, the other one happens.
- This gives us a formula for conditional probability:

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

Example:
What is the probability of A (“doubles”) given B (sum of 2 dice is 4)?

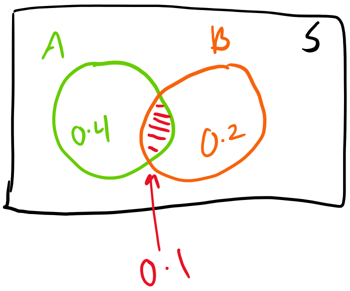
A: For a double to occur, A (“doubles”) has to occur and B (sum of 2 dice is 4) has to occur.

$$P(A \cap B) = \frac{1}{36}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{36}}{\frac{3}{36}} = \frac{1}{3}$$

- Notice that this is different from the unconditional probability $P(A) = \frac{6}{36} = \frac{1}{6}$.
- If you use Venn diagrams to calculate $P(A|B)$, $\overline{B}$ is discarded and B becomes the new sample space.

$$P(A|B) = \frac{0.1}{0.1 + 0.2} = \frac{1}{3}$$



Probability Rules

1) $P(S) = 1, P(\emptyset) = 0$
2) $P(E) \geq 0$ for any event $E \subseteq S$
3) If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$
4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5) $P(\overline{A}) = 1 - P(A)$
6) $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$

- A **two-way table** presents probabilities of all possible intersections.

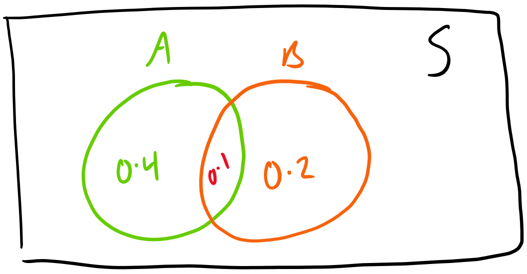
	B	$\overline{B}$	TOTAL
A	$P(A \cap B)$	$P(A \cap \overline{B})$	$P(A)$
$\overline{A}$	$P(\overline{A} \cap B)$	$P(\overline{A} \cap \overline{B})$	$P(\overline{A})$
TOTAL	$P(B)$	$P(\overline{B})$	1

Example:
Using the previous Venn diagram:

	B	$\overline{B}$	TOTAL
A	0.1	0.4	0.5
$\overline{A}$	0.2	0.3	0.5
TOTAL	0.3	0.7	1

- To find conditional probabilities using a two-way table, divide the intersection value by the row or column total.

	B	$\overline{B}$	TOTAL
A	0.1	0.4	0.5
$\overline{A}$	0.2	0.3	0.5
TOTAL	0.3	0.7	1



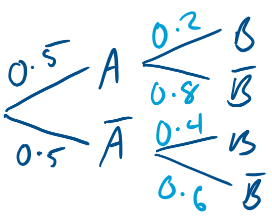
$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.1}{0.5} = \frac{1}{5}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.1}{0.3} = \frac{1}{3}$$

Tree Diagrams

- Conditional probabilities correspond to second-level (or higher) branches in a tree diagram.
- Multiply probabilities of all branches along a path to find it's probability.
- Add probabilities of all paths leading to an event to find it's probability.

Example:
Based on the previous two-way table:



$(0.5)(0.2) = 0.1 = P(A \cap B) \rightarrow P(A \cap B) = P(A)P(B|A) = 0.1$

$= (0.5) \cdot P(B|A) = 0.1$

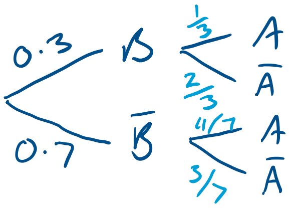
$= P(B) = \frac{0.1}{0.5} = 0.2$

$(0.5)(0.8) = 0.4 = P(\bar{A} \cap B)$

$(0.5)(0.4) = 0.2 = P(\bar{A} \cap B)$

$(0.5)(0.6) = 0.3 = P(\bar{A} \cap \bar{B})$

Exercise:
Same table, branch from B first.



$\rightarrow P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{0.1}{0.3} = \frac{1}{3}$

$\rightarrow P(\bar{A}|B) = 1 - P(A|B) = \frac{2}{3}$

$\rightarrow P(A|\bar{B}) = \frac{P(A\cap \bar{B})}{P(\bar{B})} = \frac{0.4}{0.7} = \frac{4}{7}$

$\rightarrow P(\bar{A}|\bar{B}) = 1 - P(A|\bar{B}) = \frac{3}{7}$

Law of Total Probability

- $P(A)$ can be found by decomposing A into disjoint pieces.
- Then using the sum and product rules:

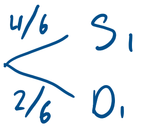
$$
\begin{aligned}
P(A) &= P(A \cap B) + P(A \cap \overline{B}) \\
&= P(B)P(A|B) + P(\overline{B})P(A|\overline{B})
\end{aligned}
$$

Example:
For the previous example,

$$P(A) = (0.3) \cdot \frac{1}{3} + (0.7) \cdot \frac{4}{7} = 0.5$$

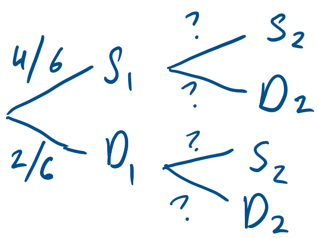
Exercise:
2 items are randomly selected without replacement from a batch of 6. The batch contains 2 defective items. Let S_i denote the event that item i inspected is satisfactory. Let D_i denote the event that item i inspected is defective. What is the probability that at least one defective item is found?

Step 1: Inspect the first item.



One item selected that is satisfactory, so 3/5 choices

Step 2: Given step 1, inspect the second item.



Intersections: $P(S_1 \cap S_2) = \frac{4}{6} \cdot \frac{3}{5} = \frac{2}{5}$

$P(S_1 \cap D_2) = \frac{4}{6} \cdot \frac{2}{5} = \frac{4}{15}$

$P(D_1 \cap S_2) = \frac{2}{6} \cdot \frac{4}{5} = \frac{4}{15}$

$P(D_1 \cap D_2) = \frac{2}{6} \cdot \frac{1}{5} = \frac{1}{15}$

$P(\text{At least one defective}) = P(D_1 \cap D_2) + P(S_1 \cap D_2) + P(D_1 \cap S_2) = \frac{1}{15} + \frac{4}{15} + \frac{4}{15} = \frac{3}{5}$

OR $1 - P(\text{No defectives}) = 1 - P(S_1 \cap S_2) = \frac{3}{5}$

Independence

- If the probability that A occurs is not affected by whether or not B occurs. i.e. $P(A|B) = P(A)$, we say that A and B are **independent.**
- We have that $P(A|B) = \frac{P(A \cap B)}{P(B)}$, so A and B are independent IFF:

$$P(A \cap B) = P(A)P(B)$$

Examples:
- Successive coin tosses are not affected by previous results, so the results of different tosses are independent.
- The events “drug present” and “positive test result are not independent, as a drug test is much more likely to be positive if the drug is present.

Exercise:
Events A and B are independent, $P(A) = 0.4, P(B) = 0.5$. Construct a two-way table and a tree diagram.

Start with what you know, and use $P(A \cap B) = P(A)P(B)$.

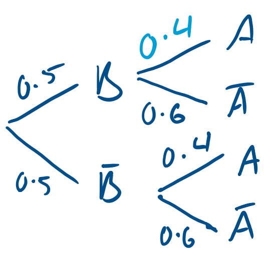
	$\boldsymbol{B}$	$\boldsymbol{\overline{B}}$	**TOTAL**
$\boldsymbol{A}$	0.2	0.2	0.4
$\boldsymbol{\overline{A}}$	0.3	0.3	0.6
TOTAL	0.5	0.5	1

$P(A \cap B) = (0.4)(0.5) = 0.2$

$P(\bar{A} \cap B) = P(B) - P(A) = 0.5 - 0.2 = 0.3$

$P(A \cap \bar{B}) = P(A) - P(B) = 0.4 - 0.2 = 0.2$

...



$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.5} = 0.4$

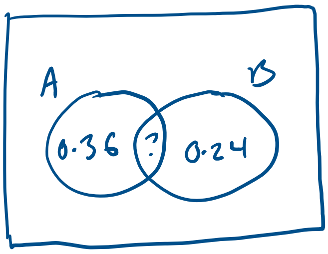
...

Exercise:

If $P(A \cap \overline{B}) = 0.36, P(\overline{A} \cap B) = 0.24, P(A|B) = 0.5$. Then A and B are:

a) Disjoint and independent.
b) Disjoint and not independent.
c) Independent and not disjoint.
d) Not independent and not disjoint.

A:



Recall that disjoint means $P(A \cap B) = 0$

$$P(B) = P(\bar{A} \cap B) + P(A \cap B)$$
$$= 0.24 + P(A \cap B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow 0.5 = \frac{P(A \cap B)}{P(A \cap B) + 0.24}$$

$$\Rightarrow 0.5(P(A \cap B) + 0.24) = P(A \cap B)$$
$$\Rightarrow 0.5(P(A \cap B)) + 0.12 = P(A \cap B)$$
$$\Rightarrow 0.12 = P(A \cap B) - 0.5(P(A \cap B))$$
$$\Rightarrow 0.12 = P(A \cap B)(1 - 0.5)$$
$$\Rightarrow P(A \cap B) = 0.24 \neq 0$$

$\therefore$ A and B are not disjoint

$P(A) = 0.36 + 0.24 = 0.6$; $P(B) = 0.24 + 0.24 = 0.48$

Recall that independence means $P(A \cap B) = P(A)P(B)$

$P(A)P(B) = (0.6)(0.48) = 0.288 \neq P(A \cap B)$

$\therefore$ A and B are not independent

Bayes' Rule

- For events A and B, Bayes' rule provides a way to reverse the order of conditional probabilities.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})}$$

- This comes directly from the definition of conditional probability, the product rule (numerator) and the law of total probability (denominator).

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

- In terms of a tree, the numerator defines one path, the denominator is the sum of paths that lead to A.

Exercise:

A drug test has 0.96 chance of positive result if the drug is present, 0.93 chance of negative result if the drug is present, 0.93 chance of negative result if the drug is not present. The unconditional probability of the drug being present is 0.007. Given a positive result, what is the probability that the drug is present?

Let A = "positive test result", B = "drug is present"

$p(A|B) = 0.96;\quad p(\bar{A}|\bar{B}) = 0.93;\quad p(B) = 0.007.$

$$p(B|A) = \frac{p(A|B)p(B)}{p(A|B)p(B) + p(A|\bar{B})p(\bar{B})}$$

$$= \frac{(0.96)(0.007)}{(0.96)(0.007) + (1-0.93)(1-0.007)}$$

$$= 0.08815$$

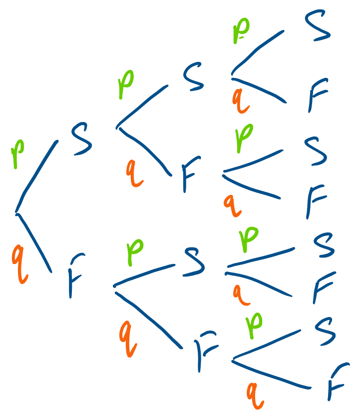
Positive test result is 91% likely to be false!

Binomial Scenario

- Fixed number of independent trials.
- 2 possible outcomes, “success” and “failure.”
- Constant probability of success for each trial.
- The quantity of interest is the total number of success.
- Notation:
  - $n = number\ of\ indenpendent\ trials$
  - $p = probability\ of\ success\ for\ a\ single\ trial.\ 0 < p < 1$
  - $q = 1 - p = probability\ of\ failure$
  - $x = number\ of\ successes$

For small n, a tree diagram can be used to work out probabilities:

x	0	1	2	3	Total
Prob.	q^3	$3pq^2$	$3p^2q$	p^3	1



- Recall the Binomial Coefficient, the number of ways to select k objects out of n (order is not important) is:

$$\binom{n}{k} = C_k^n = \frac{n1}{k1(n-k)!}$$

Example:
$\binom{3}{2} = \frac{3!}{2!(3-2)!} = 3$, so there are 3 different ways of choosing 2 items from a set of 3 times.

- An alternative interpretation is that there are $\binom{n}{x}$ ways of arranging n objects, x of one type (success) and $(n-x)$ of another type (failure): $\binom{3}{2}$ → SSF, SFS, FSS.
- Since the binomial scenario events are independent, the probability of x successes and $(n-x)$ failures in n trials (single path) is:

$$\underbrace{p \cdot p \cdot \ldots \cdot p}_{x \text{ times}} \; \underbrace{q \cdot q \cdot \ldots \cdot q}_{(n-x) \text{ times}} = p^x q^{n-x}$$

- The number of such path is $C_x^n = \binom{n}{x}$
- So the probability of x success is:

$$\binom{n}{x} p^x q^{n-x}$$

- NOTE: that the sum of all binomial probabilities is 1, as it must be.
- We see this by the binomial expansion theorem.

$$\binom{n}{0} q^n + \binom{n}{1} pq^{n-1} + \binom{n}{2} p^2 q^{n-2} + \cdots + \binom{n}{n} p^n$$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (q+p)^n = (1-p+p)^n = 1$$

Example:

The probability that an email delivered to a certain account is junk is 0.25, independently of all other messages. What is the probability that exactly 5 out of the 20 most recent messages are junk?

A: $n = 20$ is far too large for a tree diagram, so use the binomial probability formula with:

$$n = 20, x = 5, p = 0.25$$

$$P(5) = \binom{20}{5} 0.25^5 0.75^{20-5} = 0.2023$$