Probability
And
Probability Distributions

Probability and statistics are related in an important way. Probability is used as a tool. It allows us to evaluate the reliability of our conclusions about the population when we have only sample information.

Probability* *is used to quantify the likelihood, or chance, that an outcome of a random experiment will occur.

Events and the Sample Space

Data are obtained by observing either

- **uncontrolled events in nature or**
- **controlled situations in a laboratory.**

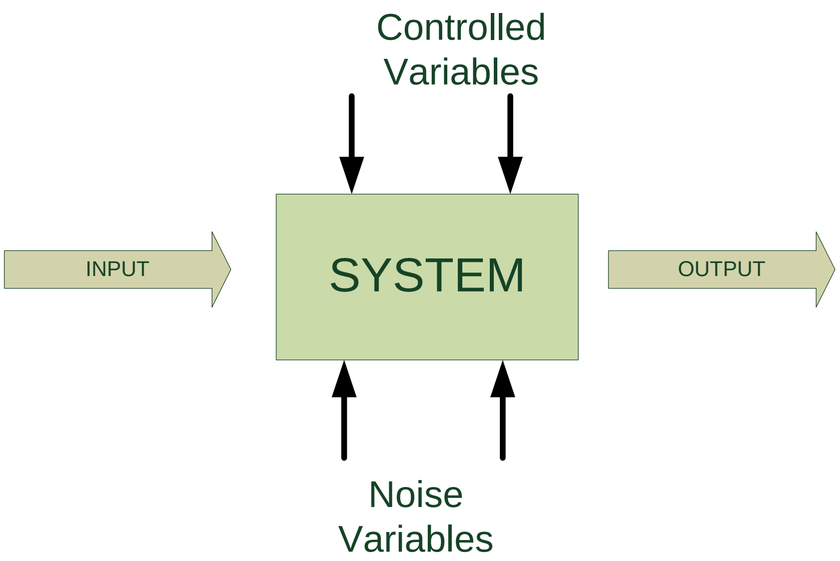
Definition: ***An experiment is the process by which an observation (or measurement) is obtained.***

The observation or measurement generated by an experiment *may or may not produce a numerical value*. Here are some examples of experiments:

- **Recording a test grade,**
- **Testing a printed circuit board to determine whether it is defective product or an acceptable product,**
- **Tossing a coin and observing the face that appears,**
- **Measuring daily rainfall.**

Definition: An experiment can result in different outcomes, even though it is repeated in the same manner every time, is called a random experiment.

Because of the uncontrollable inputs, the same settings for the controllable inputs do not result in identical outputs every time the system is measured.



Noise variables affect the transformation of inputs to outputs.

When an experiment is performed what we observe is an outcome called a simple event, often denoted by the capital E with a subscript.

Definition: ***A simple event is the outcome that is observed on a single repetition of the experiment.***

Example: Toss a die and observe the number that appears on the upper face. List the simple events in the experiment.

Six possible outcomes

Event E_1 : Observe a 1
Event E_2 : Observe a 2
Event E_3 : Observe a 3
Event E_4 : Observe a 4
Event E_5: Observe a 5
Event E_6: Observe a 6

We can now define an event as a collection of simple events.

Definition: ***An event is a collection of simple events.***

Example: We can define the events A and B for the die tossing experiment:

A: Observe an odd number

B: Observe a number less than 4.

$A=\{E_1,E_3,E_5\}$ $B=\{E_1,E_2,E_3\}$

Sample Space:

Definition: ***The set of all simple events is called the sample space S.***

Alternative Definition: ***An event is a subset of the sample space of a random experiment..***

For the die tossing experiment the sample space is

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

Toss a single coin and observe the result.

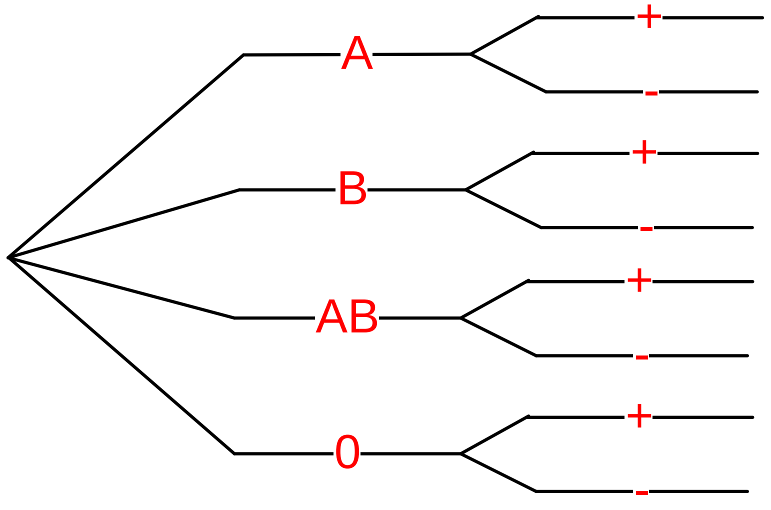
$$S = \{E_1, E_2\} = \{Tail, Head\} = \{T, H\}$$

Record a person's blood type.

$$S = \{E_1, E_2, E_3, E_4\} = \{A, B, AB, 0)$$

Or

$$S = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7\ E_8\}$$
$$= \{A+, A-, B+, B-, AB+, AB-, 0+, 0-\}$$



Tree diagram for Blood Type example

An alternative way to display the simple events is to use a *probability table*.

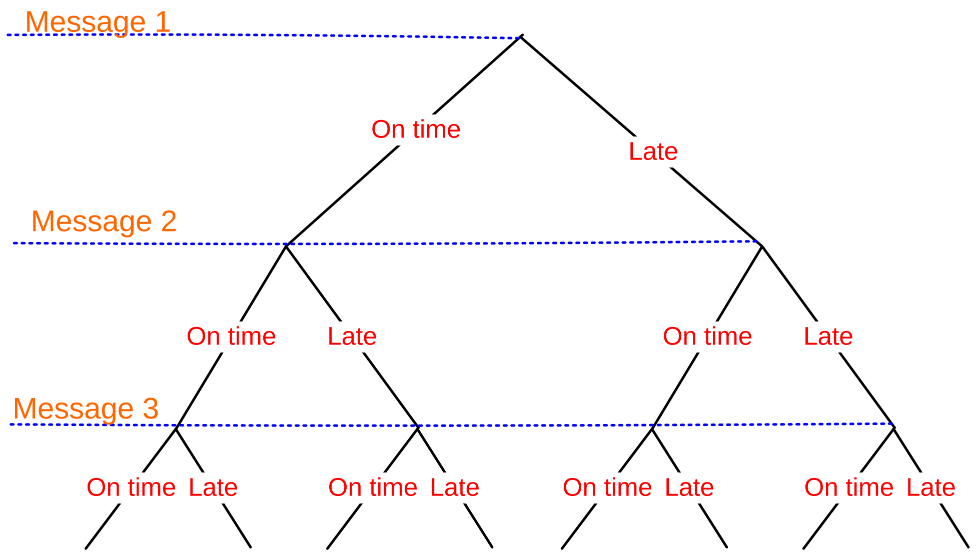
Rh Factor	Blood Type			
	A	B	AB	0
Positive	**A+**	**B+**	**AB+**	**0+**
Negative	**A-**	**B-**	**AB-**	**0-**

If the objective of the analysis is to consider only whether a particular specification is low, medium, or high, the sample space might be taken to be the set of three outcomes:

S={Low, Medium, High}

An answer to a specific question

S={yes, no}



Tree Diagram for three messages (2^3 events.)

Definition: ***A sample space is*** ***discrete*** ***if it consists of a finite or countable infinite set of outcomes.***

Definition: ***A sample space is*** ***continuous*** ***if it consists an interval (either finite or infinite) of real numbers.***

$$S=\{x \,/\, 50 < x < 100\}$$

$$S=\{x \,/\, -\infty < x < \infty\}$$

Mutually Exclusive Events

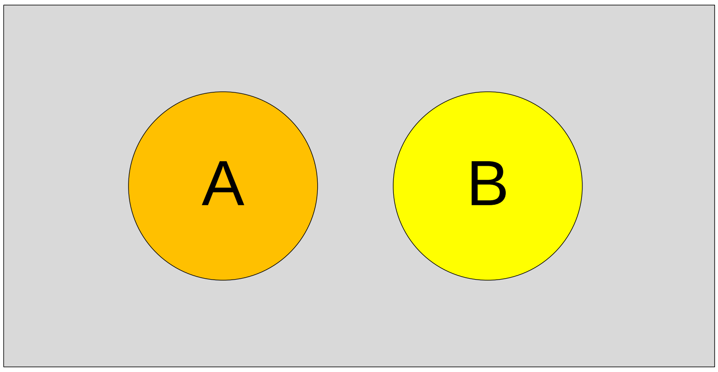
Sometimes when one event occurs, it means that another event cannot.

Two events, denoted E_1 and E_2, such that

$$E_1 \cap E_2 = \varnothing$$

are said to be mutually exclusive.

Definition:* *Two events are mutually exclusive if, one event occurs, the others cannot, and vice versa.



Venn diagram

The six simple events $\mathbf{E_1, E_2, \ldots, E_6}$ form a set of **all mutually exclusive outcomes** of the experiment.

Example: **We can define the events A and B for the die tossing experiment:**

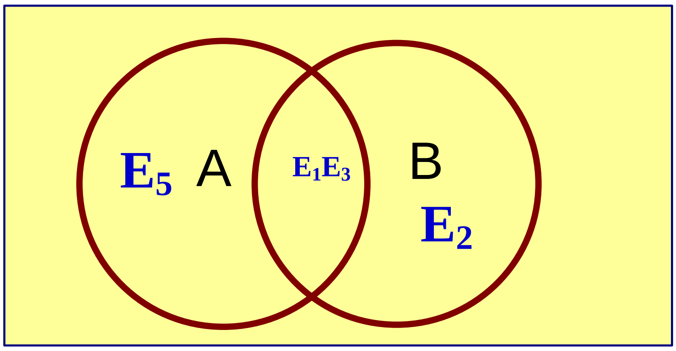
A: Observe an odd number

B: Observe a number less than 4.

$A=\{E_1,E_3,E_5\}$ $B=\{E_1,E_2,E_3\}$

In the above die-tossing experiment, events A and B are **not mutually exclusive**, because they have two outcomes in common.

$A=\{E_1,E_3,E_5\}$ $B=\{E_1,E_2,E_3\}$



Calculating probabilities using simple events

Definition: ***If an event can occur in N mutually exclusive and equally likely ways, and if m of this possess a characteristic, E, the probability of the occurrence of E is equal to m/N.***

We read P (E) as "the probability of E" and express the above definition as

$$\Pr ob(E) = P(E) = \frac{m}{N}$$

Definition: ***If an experiment is performed n times, then the relative frequency of a particular occurrence-say,-A- is***

$$\text{Re}\,lative\ \ frequency = \frac{Frequency}{n} = \frac{k}{n}$$

If we let n, the number of repetitions of the experiment, become larger and large, we will eventually generate the entire population. In this population, the relative frequency of the event A is defined the **probability of event A,** that is,

$$P(A) = \lim_{n\to\infty} \frac{Frequency}{n}$$

If we tossed a balanced, six-sided die an infinite number of times, we would expect the relative frequency for any of the six values, x=1, 2, 3, 4, 5, 6, to be 1/6.

Elementary Properties of Probability

- **Each probability must lie 0 and 1.**

$$0 < P(E) < 1$$

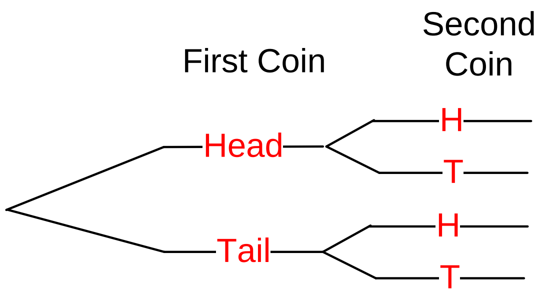
- **The sum of the probabilities for all simple events in sample space S equals 1.**

$$P(E_1) + P(E_2) + ... + P(E_N) = 1$$

Definition: ***The probability of an event A is equal to the <u>sum of the probabilities</u> of the simple events contained in A.***

Example: **Toss two fair coins and record the outcome. Find the probability of observing exactly one head in the two tosses.**

$E_1=(HH)$ $E_2=(HT)$ $E_3=(TH)$ $E_4(TT)$



Event	First Coin	Second Coin	
E_1	H	H	¼
E_2	H	T	¼
E_3	T	H	¼
E_4	T	T	¼

$$P(A) = P(E_2)+P(E_3)=1/2$$

Example: **The proportion of blood phenotypes A, B, AB and 0 in the population are reported 0.41, 0.10, 0.04 and 0.45 respectively. If a single person is chosen randomly from the population, what is the probability that he or she will have either type A or type AB blood?**

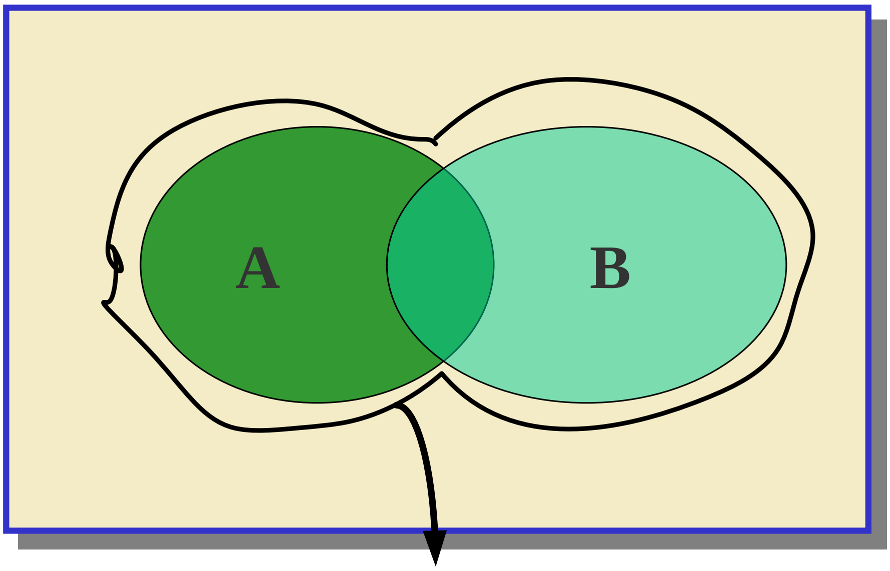
$P(A)=0.41$ $P(B)=0.10$ $P(AB)=0.04$ $P(0)=0.45$

$P(A \text{ or } AB)=P(A)+P(AB)= 0.41+0.04=0.45$

Event Relations and Probability Rules

Definition: ***The union of events A and B, denoted by***

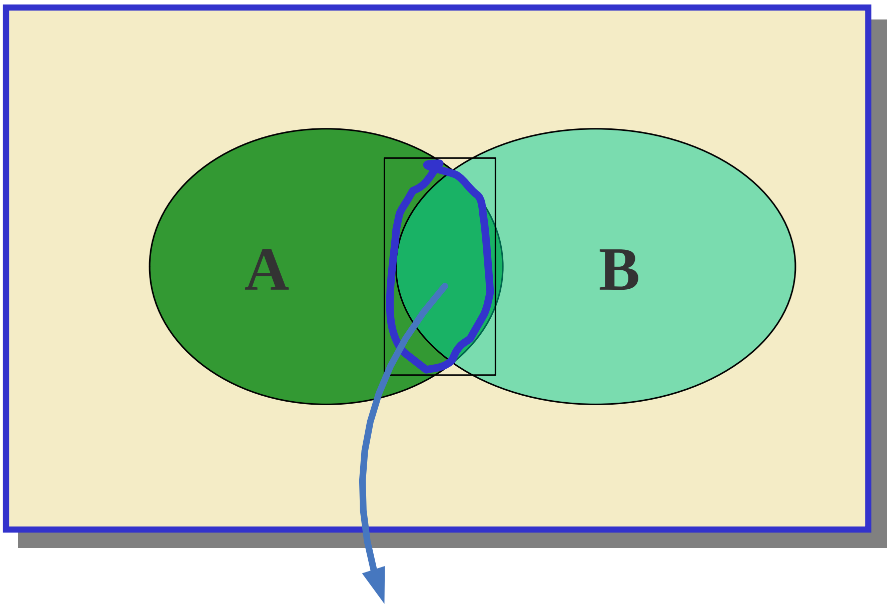
$A \cup B$ ***is the event that either A or B or both occur.***



$$A \cup B$$

Definition: ***The intersection of events A and B, denoted by***

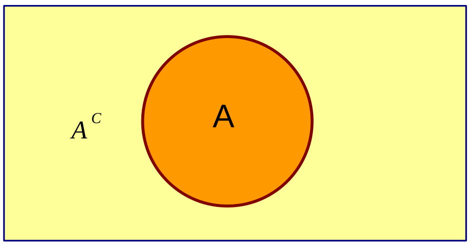
$A \cap B$ ***is the event that both A and B occur.***



$A \cap B$

Definition: ***The complement of an event A and B, denoted by***

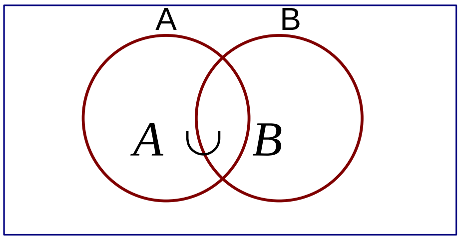
A^C ***is the event that A does not occur.***



Calculating Probabilities for Unions and Complements

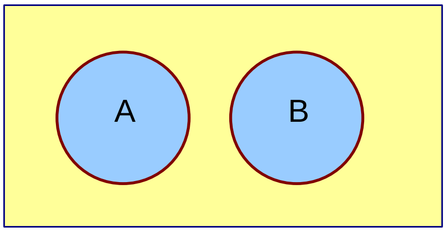
THE ADDITION RULE: ***Given two events A and B, the probability of their union,*** $A \cup B$***, is equal to***

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



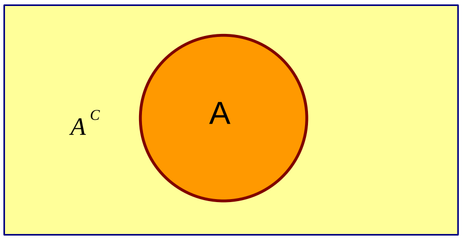
When two events A and B are mutually exclusive, then the probability that they both occur $P(A \cap B) = 0$ ***and the Addition Rule simplifies to***

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



RULE FOR COMPLEMENT:

$$P(A^C) = 1 - P(A)$$



Example: **The following Table lists history of 950 wafers in a semiconductor manufacturing process.**

Contamination	Location is Sputtering Tool		Total
	Center	Edge	
Low	514	68	582
High	112	246	358
Total	626	314	940

Suppose one wafer is selected at random.

Let H denote the event that the wafer contains high levels of contamination.

Then

$$P(H) = 358/940$$

Let C denote the event that the wafer is in the center of the sputtering tool.

Then

$$P(C) = 626/940$$

Also $P(H \cap C)$ **is the probability that the wafer is in the center of the sputtering tool and contains high levels of contamination.**

Therefore,

$$P(H \cap C) = 112/940$$

The event $H \cup C$ is the event that a wafer is from the center of the sputtering tool or contains high levels of contamination(or both).

From the Table,

$$P(H \cup C) = P(H) + P(C) - P(H \cap C)$$
$$= 358/940 + 626/940 - 112/940 = 872/940$$

```
Tossing two dies and finding the probability one of them is 1.
MTB > random 1000 c1 c2;
SUBC> integer 1 6.
MTB > let c3=c1=1 or c2=1
MTB > let k1=sum(c3)/count(c3)
MTB > print k1
 Data Display

K1     0.315000

MTB > let k2=11/36
MTB > print k2
 Data Display
K2     0.305556
MTB > random 10000 c4 c5;
SUBC> integer 1 6.
MTB > let c6=c4=1 or c5=1
MTB > let k1=sum(c6)/count(c6)
MTB > print k1
 Data Display
K1     0.299700
MTB > random 100000 c7 c8;
SUBC> integer 1 6.
MTB > let c9=c7=1 or c8=1.
MTB > let k1=sum(c9)/count(c9)
MTB > print k1

Data Display

K1     0.305690
real value is 1/6+1/6-1/6*1/6=2/6-1/36=11/36=0.305555
```

INDEPENDENCE,
CONDITIONAL PROBABILITY,
AND
THE MULTIPLICATION RULE

Definition: ***Two events, A and B are said to be independent if and only if the probability of event B is not influenced or changed by the occurrence of event A, or vice versa.***

Alternatively two events A and B are independent if

$$P(A \setminus \mathrm{B}) = \mathrm{P(A)} \quad \textit{or} \quad P(B \setminus \mathrm{A}) = \mathrm{P(B)}.$$

The probability of an event A, given that the event B has occurred, is called conditional probability of A given B has occurred, denoted by $P(A \setminus \mathrm{B})$.

Tossing Dice Consider tossing a single die two times, and define two events:

A: Observe a 2 on the first toss

B: Observe a 2 on the second toss

If the die is fair, the probability of event A is P(A)=1/6.

Regardless of whether event A has or has not occurred, the probability of observing a 2 on the second toss is still 1/6. We could write

P(B given that A occurred)=1/6

P(B given that A did not occur)=1/6.

Since the probability of event B is not changed by the occurrence of event A, **we say that A and B are independent events.**

THE GENERAL MULTIPLICATION RULE

The probability that both A and B occur when the experiment is performed is

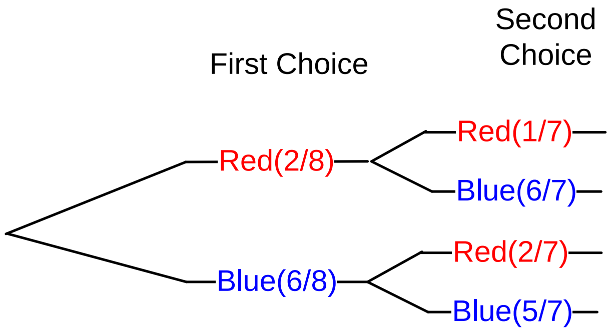
$$P(A \cap B) = P(A)P(B \setminus \mathrm{A})$$ ***or***

$$P(A \cap B) = P(B)P(A \setminus \mathrm{B})$$

Example: Eight toys are placed in a container. The toys identical except for color; **two** are **red** and **six** are **blue**. **What is the probability that the child chooses the two red toys?**

R: red toy is chosen **B: Blue Boy is chosen**

A=(R on the first choice) $\cap$ (R on the second choice)



P(A)= (2/8)(1/7)=2/56=1/28

CONDITIONAL PROBABILITIES

The conditional probability of event A, given that event B has occurred is

$$P(A\setminus \mathrm{B})=\frac{\mathrm{P(A\cap B)}}{\mathrm{P(B)}}\quad if\quad \mathrm{P(B)}\neq 0$$

The conditional probability of event B, given that event A has occurred is

$$P(B\setminus \mathrm{A})=\frac{\mathrm{P(A\cap B)}}{\mathrm{P(A)}}\quad if\quad \mathrm{P(A)}\neq 0$$

MULTIPLICATION RULE FOR INDEPENDENT EVENTS

If two events A and B are independent, the probability that both A and B occur is

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

Similarly, if A, B and C are mutually independent events, then the probability that A, B and C all occur is

$$P(A\cap B\cap C)=P(A)P(B)\mathrm{P(C)}$$

Example: Suppose that in the general population, there are 51% men and 49% women, and that the proportion of colorblind men and women are shown in the probability table below:

	Men (B)	Women(B^C)	Total
Colorblind(A)	**0.04**	**0.002**	**0.042**
Not Colorblind(A^C)	**0.47**	**0.488**	**0.958**
Total	**0.51**	**0.49**	

If a person is drawn at random from this population and is found to be a man (**event B**) **what is the probability that man is colorblind** (**event A**)?

If we know that the event B has occurred, **we must restrict our focus** to only the 51% of the population that is male.

$$P(A \setminus B) = \frac{P(A \cap B)}{P(B)} = \frac{0.04}{0.51} = 0.078$$

The probability of being colorblind, given that the person is male, is 4% of the 51%.

What is the probability of being colorblind, given that the person is female?

Now we are restricted to only the 49% of the population that is female, and

$$P(A \setminus B^C) = \frac{P(A \cap B^C)}{P(B^C)} = \frac{0.002}{0.49} = 0.004$$

Notice that the probability of event A changed, depending on whether event B occurred.

This indicates that these two events are dependent.

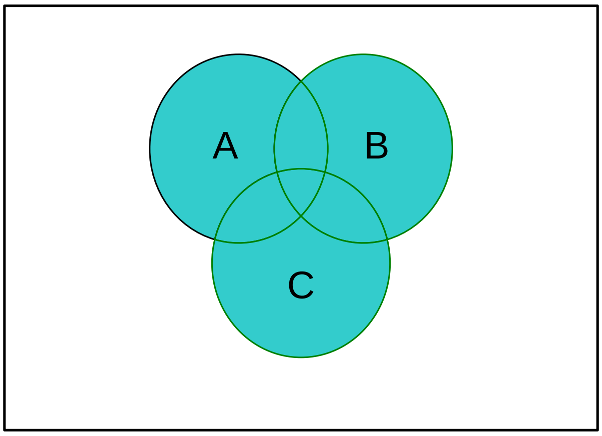
What is the difference between Mutually Exclusive and Independent Events?

- **When two events are mutually exclusive or disjoint, they cannot both happen when the experiment is performed. Once the event B has occurred, event A cannot occur, so that P(A\B)=0, or vice versa.(A=Die score 1 ,B=Die score 2). The occurrence of event B certainly affects the that event A can occur.**
- **Therefore, mutually exclusive events must be dependent.**
- **When two events are mutually exclusive events or disjoint, $P(A\cap B)=0$ and $P(A\cup B)=P(A)+P(B)$**
- **When two events are independent. $P(A\cup B)=P(A)+P(B)- P(A\cap B)$**

Example: Use relationship above to fill in the table below.

P(A)	P(B)	Conditions for A and B	$P(A\cap B)$	$P(A\cup B)$	P(A\B)
0.3	0.4	Mutually Exclusive	0	0.7	0
0.3	0.4	Independent	0.12	0.58	0.3
0.1	0.5	Mutually exclusive and dependent	0	0.6	0
0.2	0.5	Independent	0.10	0.6	0.2

Three or More Events



P(A∪B∪C)=P(A∪B)∪C)=P(A∪B)+P(C)-P(A∪B)∩C)

P(A∪B∪C)=P(A)+P(B)-P(A∩B)+P(C)-

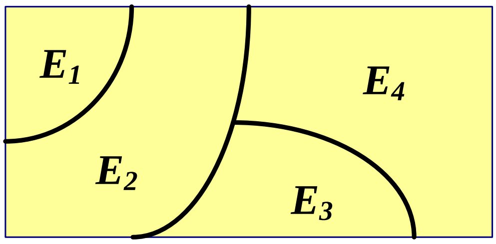
[P(A∩C)+P(B∩C)-P(A∩B∩C)]

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

Mutually Exclusive Events

A collection of events E_1, E_2,…,E_k is said to be mutually exclusive events if for all pairs

$$E_i \cap E_j = \varnothing.$$



EXAMPLES:

Example: Suppose that we have a fuse box containing **20 fuses**, of which **5 are defective**. If **2 fuses** are selected at random and removed from the box in succession without replacing the first, **what is the probability that both fuses are defective**?

A: The event that first fuse is defective

B: The event that second fuse is defective

$$P(A \cap B)=(5/20)(4/19)=(1/4)(4/19)=1/19$$

Example: One bag contains 4 white balls and 3 black balls, and a second bag contains 3 white balls and 5 black balls. One ball is drawn from the first bag and placed unseen in the second bag. **What is the probability that a ball now drawn from the second bag is black?**

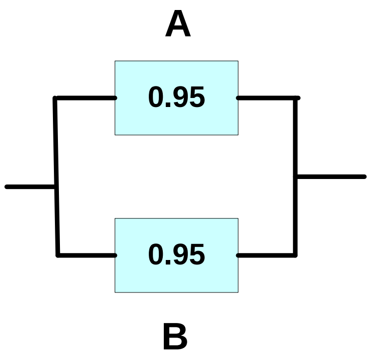
Let B_1, B_2, ad W_1 represent, respectively, the drawing of black ball from bag1, a black ball from bag 2, and a white ball from bag 1.

$$P[(B_1 \cap B_2) \text{ or } (W_1 \cap B_2)] = P(B_1 \cap B_2) + P(W_1 \cap B_2)$$

$$= P(B_1)\, P(B_2 \backslash B_1) + P(W_1)\, P(B_2 / W_1)$$

$$= (3/7)(6/9) + (4/7)(5/9) = 38/63$$

Example: The following circuit operates only if there is a path of functional devices from left to right. Assume that the devices fail independently. **What is the probability that the circuit operates?**



There is a path if at least one device operates. The probability that the circuit operates is

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

$$= (0.95)+(0.95)-(0.95)(0.95)= 0.9975$$

Alternative Solution:

Intersection of complements

$$P(A^c\cap B^c)= 1-P(A\cup B)$$

Or $P(A\cup B)=1- P(A^c\cap B^c)$

Because of the events A and B independent

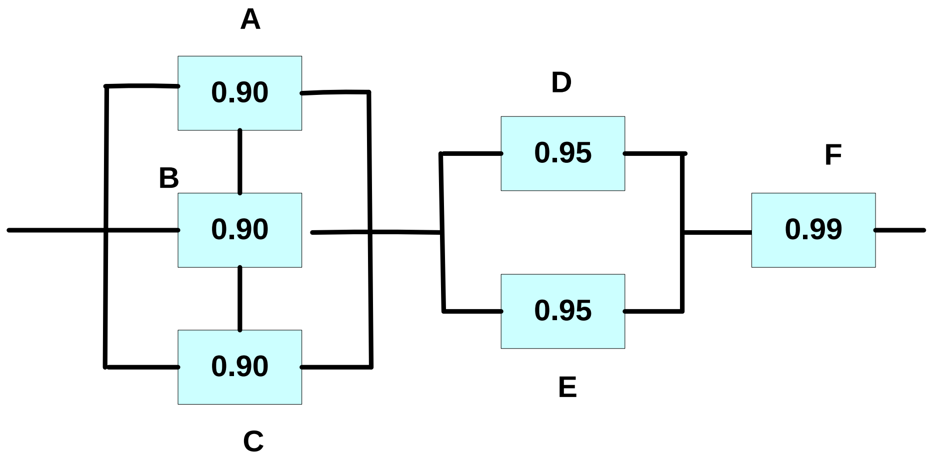
$$P(A\cup B)=1- [P(A^c)P(B^c)]=1-(1-0.95)(1-0.95)$$

$$=1-(0.05)(0.05)=1-0.0025 = \mathbf{0.9975}$$

De Morgan Laws

- ***The complement of the intersection of any number of sets equals the union of their complements.***
- ***The complement of the union of any number of sets equals the intersection of their complements.***

Example: The following circuit operates only if there is a path of functional from left to right. Assume that the devices fail independently. **What is the probability that the circuit operates?**



$P[(A\cup B\cup C)\cap(D\cup E)\cap(F)]=?$

Using complements

$P[(A\cup B\cup C)\cap(D\cup E)\cap(F)]$

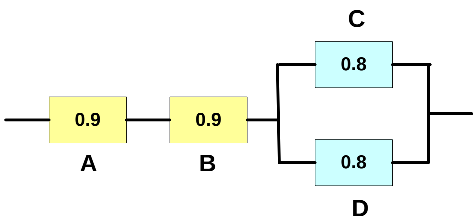
$= \{[1-P(A^c)P(B^c)P(C^c)][1-P(D^c)P(E^c)][P(F)]\}$

$=\{[1-(1-0.90)^3][1-(1-0.95)^2](0.99)\}=\mathbf{0.987}$

Example: An electrical system consists of four components. The system works if components A and B work and either of the components C or D work. The probability of working (reliability) of each component is shown in the Figure. Find the probability that

(a) **The entire system works** and

(b) **The component C does not work given that the entire system works**. Assume that four components work independently.



In this configuration of the system A, B, and the subsystem C and D constitute a serial circuit system, whereas the subsystem C and D itself is a parallel circuit system.

(a) The probability that the entire system works can be calculated as the following:

$$P[A\cap B\cap(C\cup D)]=P(A)P(B)P(C\cup D)$$

$$= P(A)P(B)[1-P(C^C\cap D^C)]= P(A)P(B)[1-P(C^C)P(D^C)]$$

$$=(0.9)(0.9)[1-(1-0.8)(1-0.8)]=0.7776$$

(b) To calculate the conditional probability in this case, notice that

$$P=\frac{P(\text{the system works but C does not work})}{\text{P(the system works)}}$$

$$P=\frac{P(A\cap B\cap C^c\cap \text{D})}{\text{P(the system works)}}=\frac{(0.9)(0.9)(1-0.8)(0.8)}{0.7776}=0.1667$$