# Introduction to Probability

In everyday conversation, the term probability is a measure of one's belief in the occurrence of a future event. We accept this as a meaningful and practical interpretation of probability but seek a clearer understanding of its context, how it is measured, and how it assists in making inferences.

The necessity of considering the concept of probability is due to the existence of random or stochastic process, when observations or outcomes cannot be predicted with certainty.

For example, the blood pressure of a person at a given point in time cannot be predicted with certainty, and we never know the exact load that a bridge will endure before collapsing into a river.

There are many types of ways to measure probability

1. Classical
   - Consider a finite number of outcomes
   - Assume each outcome is equally likely
   - $P(A) = \frac{\#\, number\ of\ ways\ A\ can\ occur}{total\ \#\ of\ outcomes}$
   - Example: A and B play a fair game 3 times, what is the probability B wins exactly 2 games?

AAA  BBB
AAB  BBA
ABA  BAB
ABB  BAA.

3 of 8 => $\frac{3}{8}$ = P(B wins exactly 2 games)

2. Frequency/Counts (empirical, "based off of samples or data")
   - After observing a sample situation (of a "large" sample size) expect that the actual probability of an event to be modeled by the sample
   - "Long-run" observations
   - Example: Suppose a pharmacist stays late to work overtime 30 days out of 200 days, the probability that this pharmacist works overtime is likely to be…?

$\frac{30}{200} = 0.15$ or 15% ≅ P(of working overtime)

3. Axiomatic (Building Blocks, Ground Rules)
    - Formal
    - Define probability as a function acting on a set
    - This function must possess certain properties
    - Suppose S is a sample space associated with an experiment. To every event A in S (A is a subset of S), we assign a number, P(A), called the probability of A, so that the following axioms hold:
      Axiom 1: $P(A) \geq 0$.
      Axiom 2: $P(S) = 1$.
      Axiom 3: If A1, A2, A3, . . . form a sequence of pairwise mutually exclusive events in S (that is, $A_i \cap A_j = \emptyset$ if $i \neq j$), then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$

More on these later...

4. Subjective
    - Probability is based on personal opinion
    - Example: What do you feel is the chance that the senate will pass a reform on Healthcare this year? / Weather?

About ... 25% chance...

Prior to introducing the notion of probability, there are some terms that need to be defined. These are:

1. Random experiment: A random experiment is a process that produces one of many possible outcomes. In addition, outcomes are *unpredictable*. Does Not mean all Events are equally likely though.

2. Sample space, or (S): A sample space, denoted by S, is a set that *lists all* the outcomes produced by a random experiment. Often each outcome is referred to as a 'sample point' or an 'element'
   There can be different ways of exhibiting the sample space
    - A coin is tossed twice...

$S = \{(H,H), (H,T), (T,H), (T,T) \mid$ when $(x,y)$, $x$ is outcome from 1st toss, $y$ is outcome from 2nd toss$\}$

OR Another Representation of S could be...

$S = \{x = 0, 1, 2 \mid x$ is The number of Heads observed$\}$

⋮

- A Coin tossed until a head Appears
  $S = \{H, TH, TTH, \cdots\}$ OR $S = \{x = 1, 2, 3, \ldots \mid x$ is The number of tosses until 1st Head$\}$

"Set that is smaller to or equal to"

"A collection of possible outcomes"

3. Event: An event is a *subset* of the sample space, which consists of elements that share a common characteristic. That is, if A represents an event, then $A \subseteq S$.
   - When waiting for a bus, consider how long you could be waiting for…

     $S = \{0 \le x < \infty \mid \text{where } x \text{ is time measured in minutes}\}$

     OR

     $S = [0, \infty)$

   - Let A be the event you wait for at least 10 minutes…

     $A = \{10 \le x < \infty \mid \text{where } x \text{ is time measured in minutes}\}$

     OR

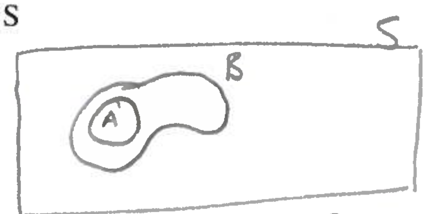
     $A = [10, \infty)$

## Set Theory and Probability

Consider two events, A and B, where $A \subseteq S$, $B \subseteq S$

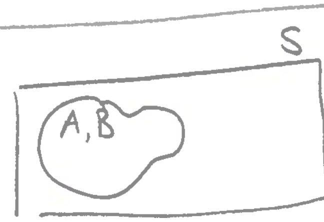
Containment/Subset:

if $A \subseteq B : \forall x \in A \Rightarrow x \in B$ eg.




Equality:

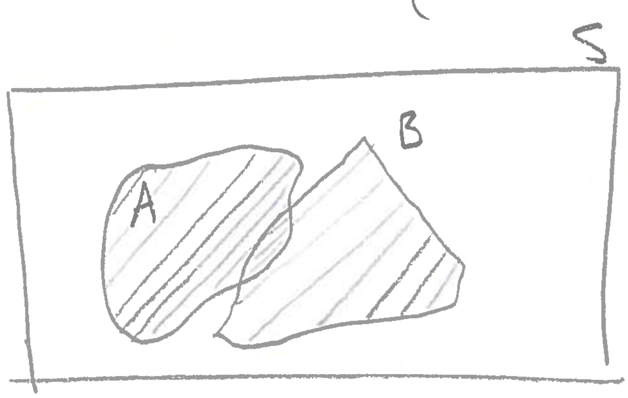
if $A = B : \forall x \in A \Rightarrow x \in B$ and $\forall x \in B \Rightarrow x \in A$ eg.




"B is another Name for A"

Union of A and B - $A \cup B$: The union of events A and B - denoted by $(A \cup B)$ - is a subset of the sample space which consists of elements found in A or elements found in B or elements found in BOTH A and B. A Venn diagram of $(A \cup B)$ is to be drawn in the space below.

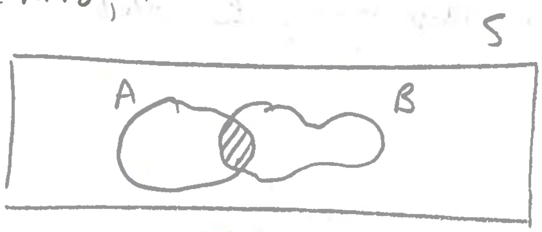
if $x \in A \cup B$ then, $\forall x$, $x \in A$ 'or' $x \in B$
(Note $x \in A$ 'and' B)




▨ $A \cup B$

**Intersection of A and B - A ∩ B:** The intersection of events A and B - denoted by (A ∩ B) - is a subset of S, which consists of elements found in BOTH A, and B. The Venn diagram is given below.

if $x \in A \cap B$, then $\forall x;\ x \in A$ and $x \in B$




▨ - $A \cap B$

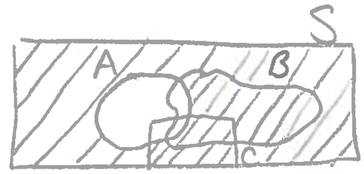
**Complement of A - $A^c$ or $\bar{A}$:** The complement of event A, $A^c$, or $\bar{A}$, is a subset of the sample space consisting of all elements in S, which are not found in A. In short, $A^c$ can be thought of as "'not A" or "everything but A". The Venn diagram of $A^c$ is given by:

$n(A \cup A^c) = n(S)$

$n(A \cap A^c) = \phi$

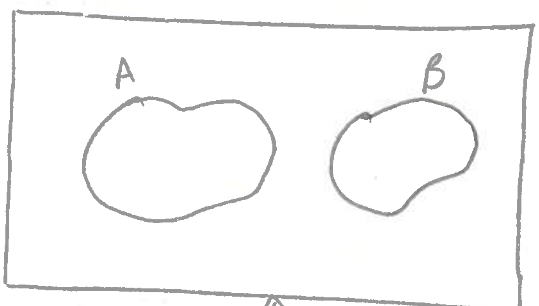
if $x \in A^c$, Then $\forall x;\ x \notin A$

eg. what is The complement of white? (Not white!)



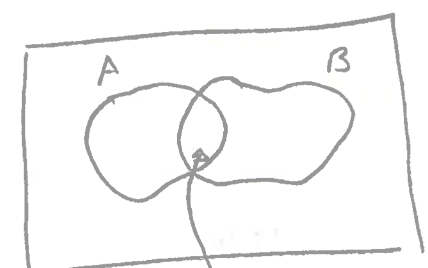

▨ - $A^c$

**Mutually Exclusive Events:** Two events, A and B, are said to be mutually exclusive when the joint intersection of A and B is an empty, or a null, set. Therefore, A and B are mutually exclusive events when (A ∩ B) = ∅. A and $A^c$ are mutually exclusive events since there are no common elements shared between A and $A^c$. (Dis joint)




"Not joined"

OR

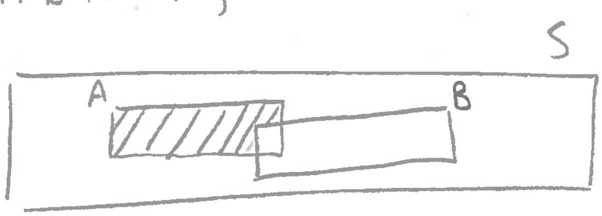



$\phi$, empty
Nothing can exist in Here.

Side note. A-B set difference

$A - B = A \cap B^c$

if $x \in A \cap B^c$ then $\forall x;\ x \in A$ and $x \notin B$




▨ - $A \cap B^c$

Recall the Axioms of probability (foundational rules)

Axiom 1: P(A) ≥ 0.

probability of Any Event must be Non-negative ⇒ 'Probability cannot be negative'

Axiom 2: P(S) = 1.

Sample space, Considers Every outcome So

$P(S) = 100\% = 1$

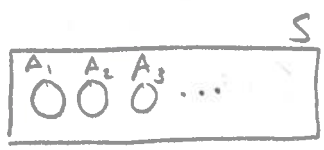
⇒ if any Event occurs, then what is The probability that an event occured? ... 100%

Axiom 3: If A1, A2, A3, . . . form a sequence of pairwise mutually exclusive events in S (that is, $A_i \cap A_j = \varnothing$ if $i \neq j$), then

"No overlap"

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$

Don't over lap.




Then $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$

Can split up Union of M.E. events into Additive Model.

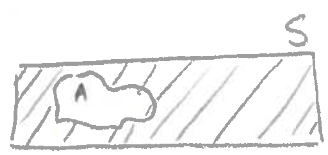
## Some Rules of Probability

From the Venn Diagrams developed the previous pages; we can derive the following probability formulae.

Law of Complement: The probability of A not occurring, or $P(A^c)$, is given by:
$P(A^c) = 1 - P(A)$.

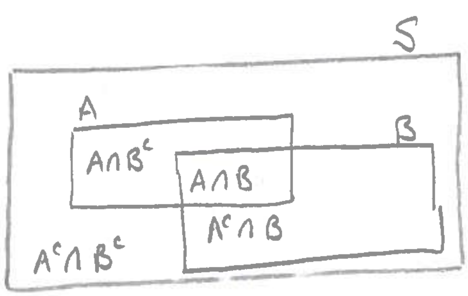
if $P(S) = 1$, $P(A) + P(A^c) = P(S) = 1$

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




▨ - $A^c$

Total Law of Probability (Simplified): The probability of A is given by:
$P(A) = P(A \cap B) + P(A \cap B^c)$

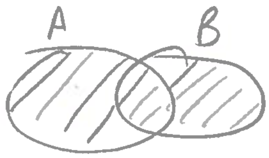



$P(A) = P(A \cap B^c) + P(A \cap B)$

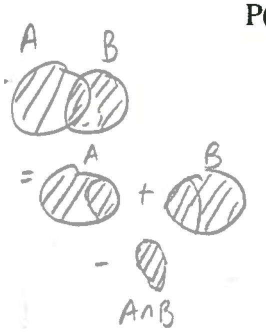
OR use Contingency tables . . .

margins or marginal probabilities

| | B | $B^c$ | |
|---|---|---|---|
| A | $P(A \cap B)$ | $P(A \cap B^c)$ | $P(A) = P(A \cap B) + P(A \cap B^c)$ |
| $A^c$ | $P(A^c \cap B)$ + | $P(A^c \cap B^c)$ | $P(A^c) = P(A^c \cap B) + P(A^c \cap B^c)$ |
| marginal Probabilities | $P(B) = P(A \cap B) \ldots$ | $P(B^c) = P(A \cap B^c) + P(A^c \cap B^c)$ | 1 or 100% |



$\boxtimes$ $A \cup B$





**Addition Law:** The probability of A or B is given by:

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

Proof:

$P(A) = P(A \cap B) + P(A \cap B^c)$ By Total law...

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

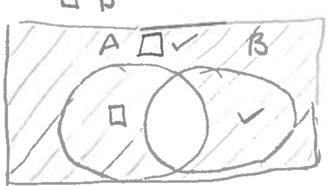
Then $P(A) + P(B) = \underbrace{P(A \cap B) + P(A \cap B^c) + P(A \cap B)}_{\text{2 of them}} + P(A^c \cap B)$

So $P(A) + P(B) - P(A \cap B) = P(A \cap B) + P(A \cap B^c) + P(A^c \cap B) = P(A \cup B)$ □

**DeMorgan's Laws:** "Bring The c inside Brackets"

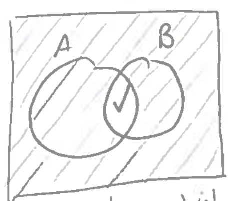
(i) $P(A^c \cup B^c) = P(A \cap B)^c$

✓ $A^c$
□ $B^c$



Any Section That has □ or ✓ or Both $A^c \cup B^c$

✓ $A \cap B$



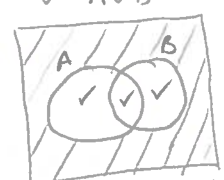
Any Section which has Nothing is $(A \cap B)^c$

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

✓ $A^c$, □ $B^c$

Any Section That has Both □, ✓ $A^c \cap B^c$

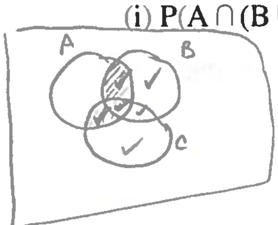
✓ $A \cup B$



Any Section with no mark is $(A \cup B)^c$

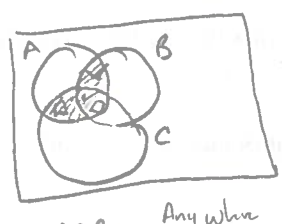
**Distributive Laws:** Let A, B, and C be events such that A, B, C ⊆ S.

(i) $P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C))$



✓ $B \cup C$

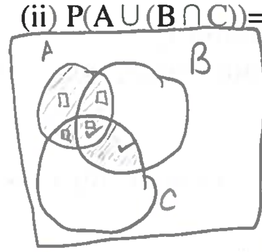
Any where ✓ intersects A is $A \cap (B \cup C)$



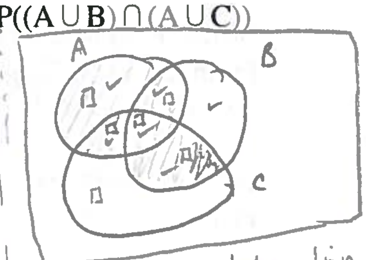
✓ $A \cap B$
□ $A \cap C$

Any where ✓ or □ or Both is Union

(ii) $P(A \cup (B \cap C)) = P((A \cup B) \cap (A \cup C))$



✓ $B \cap C$
□ A
union is Any where □ or ✓ or Both



✓ $A \cup B$
□ $A \cup C$

Intersection is any where Both □ and ✓ are

Notice that these properties can be extended to any finite collection (or countably infinite collection) of sets

- Example. 4 letters could be more...

$A \cup (B \cap C \cap D) \overset{\text{Association}}{=} A \cup (B \cap (C \cap D)) \overset{\text{Distribution}}{=} (A \cup B) \cap (A \cup (C \cap D))$

$\overset{\text{Distribution}}{=} (A \cup B) \cap (A \cup C) \cap (A \cup D)$

Examples:

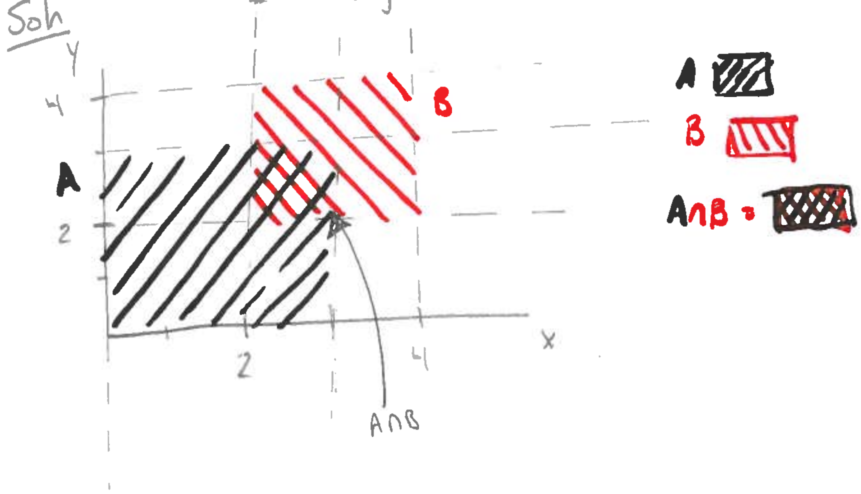
1. Use a Cartesian plane to Diagram $A \cap B$.

where $A = \{(x,y) \mid 0<x<3,\ 0<y<3\}$

$B = \{(x,y) \mid 2<x<4,\ 2<y<4\}$

<u>Soln</u>

$A \cap B = \{(x,y) \mid 2<x<3,\ 2<y<3\}$ □




2. Use a Number line to Diagram $A^c \cap B \cap C$.

where $A = \{x \mid 0 \le x \le 1\}$, $B = \{x \mid 0 \le x \le 3\}$, $C = \{x \mid -1 \le x \le 2\}$

$A^c = \{x \mid x<0 \cup x>1\}$

<u>Soln</u>


where Do The intersect?

$A^c \cap B \cap C = \{x \mid 1 < x \le 2\} = (1, 2]$ □

3. let $P(A) = 0.4$, $P(B) = 0.5$, $P(A \cap B) = 0.1$. Find the probability of A or B but not Both.

<u>Soln</u>

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

<u>Note</u>: $P(A) = P(A \cap B) + P(A \cap B^c) \Rightarrow 0.4 = 0.1 + P(A \cap B^c) \Rightarrow P(A \cap B^c) = 0.3$

$P(B) = P(A \cap B) + P(A^c \cap B) \Rightarrow 0.5 = 0.1 + P(A^c \cap B) \Rightarrow P(A^c \cap B) = 0.4$

So $P(A \cap B^c) + P(A^c \cap B) = 0.3 + 0.4 = 0.7$ □

<u>OR</u>

| | B | $B^c$ | |
|---|---|---|---|
| A | 0.1 | 0.3 | 0.4 |
| $A^c$ | 0.4 | 0.2 | 0.6 |
| | 0.5 | 0.5 | 1 |

4. if $A_1 \cup A_2 = S$, $A_1 \cap A_2 = \phi$, $P(A_1) = x$, $P(A_2) = y$.

$3x - y = \frac{1}{2}$. Find $x, y$

Sol

Since $A_1$ & $A_2$ are disjoint

$P(A_1 \cup A_2) = 1$ And $P(A_1) + P(A_2) = P(A_1 \cup A_2) = 1$

$\Rightarrow x + y = 1 \Rightarrow x = 1 - y$

$3x - y = \frac{1}{2}$

$\hookrightarrow 3(1-y) - y = \frac{1}{2} \Rightarrow 3 - 3y - y = \frac{1}{2}$

$\Rightarrow 3 - 4y = \frac{1}{2} \Rightarrow 3 - \frac{1}{2} = 4y \Rightarrow 2.5 = 4y$

$\Rightarrow \boxed{y = 0.625 = \frac{5}{8}}$

$x = 1 - y \Rightarrow x = 1 - \frac{5}{8} = \boxed{\frac{3}{8} = x}$

5. 28% of males smoke cigarettes (let A), 7% of males smoke cigars (let B), and 5% Smoke Both

a) what % Don't smoke?

Sol

Complement. Addition Rule.

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

$= 1 - [0.28 + 0.07 - 0.05] = 1 - 0.3 = 0.7$ □

OR 70% of males Do not smoke

b) What % smoke cigars not cigarettes?

Sol

$P(B \cap A^c) = P(B) - P(A \cap B) = 0.07 - 0.05 = 0.02$

2% of males smoke cigars but not cigarettes □

OR

| | $B$ | $B^c$ | |
|---|---|---|---|
| $A$ | 0.05 | 0.23 | 0.28 |
| $A^c$ | 0.02 | 0.7 | 0.72 |
| | 0.07 | 0.93 | 1 |

optional R code.

Exercise 2.1

Suppose a family contains two children of different ages, and we are interested in the gender of these children. Let F denote that a child is female and M that the child is male and let a pair such as FM denote that the older child is female and the younger is male. There are four points in the set S of possible observations:

S = {FF, FM, MF, MM}.

Let A denote the subset of possibilities containing no males; B, the subset containing two males; and C, the subset containing at least one male.
List the elements of A, B, C, A ∩B, A ∪ B, A ∩ C, A ∪ C, B ∩ C, B ∪ C, and C ∩ $B^c$.

```
#Using Software (R)
#The following computer code will answer this question, if you study it carefully you may learn how
#to apply the software in multiple situations.

library(prob)
#this includes the probability library in the program

family=function (times)
#family is the name of the function it could be changed to whatever you like
{
   temp <- list()
   for (i in 1:times) {
      temp[[i]] <- c("F", "M")
#"F", "M",... are the female and male options of each child they can be changed to other #outcomes
   }
   res <- expand.grid(temp, KEEP.OUT.ATTRS = FALSE)
   names(res) <- c(paste(rep("child", times), 1:times, sep = ""))
# "child" is the name of the column
   return(res)
}
#this sets up the sample space

S=family(2)
A=S[1,]
B=S[4,]
C=S[2:4,]
A #lists the elements in event A
B #lists the elements in event B
C #lists the elements in event C
intersect(A,B) #lists the elements in this intersection
union(A,B) #lists the elements in this union
intersect(A,C) #lists the elements in this intersection
union(A,C) #lists the elements in this union
intersect(B,C) #lists the elements in this intersection
union(B,C) #lists the elements in this union
setdiff(C,B) # lists the elements in C but not in B

#Note: The complement of A can be found by: setdiff(S,A)
```