CHAPTER 2. BASIC SET THEORY

2.1 Basic definitions

Sets are the basis on which modern probability theory is defined. A set is a well-defined collection of objects. The objects are called "elements" or "members" of the set. Typically a set is denoted by uppercase letters A, B, C, P etc. and the elements are denoted by lowercase letters a,b,c, x, y etc. A set is completely described by its members. The description can be achieved either by (i) listing (i.e., *enumerating*) the members, e.g.:

> X = {a,e,i,o,u} when describing the set of vowels.

or, (ii) by stating the *membership rule*, e.g.:

> X={x: x is an integer between 1 and 100} when describing the set of the first 100 natural numbers. The second approach is more powerful.

Symbolically, $x \in A$ states "x is an element of A," and $x \notin A$ denotes otherwise.

Universal set: In the context of a problem, all sets of interest may be subsets of some large fixed set. This superset is called the universal set.

Null set: The null set, or the empty set, is the set wih no elements. It is denoted by the special symbol $\varnothing$.

Countable set: A set is countable if its members can be placed in a one-to-one correspondence with the set of natural numbers. Otherwise, the set is *uncountable*.

2.2 Set relations

Subset: If every element of a set A is also an element of set B, then A is called a subset of B, written symbolically as: $A \subseteq B,$ or $B \supseteq A$. If A is a subset of B and B has at least one element that does not belong to A, then A is a proper subset of B, written symbolically as: $A \subset B,$ or $B \supset A$.

Superset: If $A \subseteq B$, then B is called a superset of A. If $A \subset B$, then B is a proper superset of A.

Equality of sets: If every element of A is an element of B and vice versa, i.e., $A \subseteq B$ and $B \subseteq A$, then the two sets are equal, written as: A = B.

Transitivity: If $A \subseteq B$ and $B \subseteq C,$ then $A \subseteq C.$

<u>Example</u>: Find the sets that are equal among the following:

$$A = \{x:\ x^2 - 4x + 3 = 0\}$$

$$B = \{x:\ x^2 - 3x + 2 = 0\}$$

$$C = \{x: x \in P, x < 3\}, \qquad P = \text{set of positive integers}$$

$$D = \{x: x \in P, x < 5, x \text{ is odd}\}, \qquad P = \text{set of positive integers}$$

Answer: A = D, B = C.

2.3 Operations on sets

Venn diagrams can be used for graphical representation of sets and their relations.

Operations on one or more sets produce new sets. Basic operations are: complementation, differencing, symmetric differencing, etc.

2.3.1 *Boolean combination of sets*

Given two sets *A* and *B*, their intersection *C* is the set such that it contains only those elements that belong to both:

$$C = A \cap B \Rightarrow C = \{x : x \in A \text{ and } x \in B\} \tag{2.1}$$

The union is the set *D* such that it contains elements that belong to *A* or *B* or both:

$$D = A \cup B \Rightarrow D = \{x : x \in A \text{ or } x \in B \text{ or } x \in A \cap B\} \tag{2.2}$$

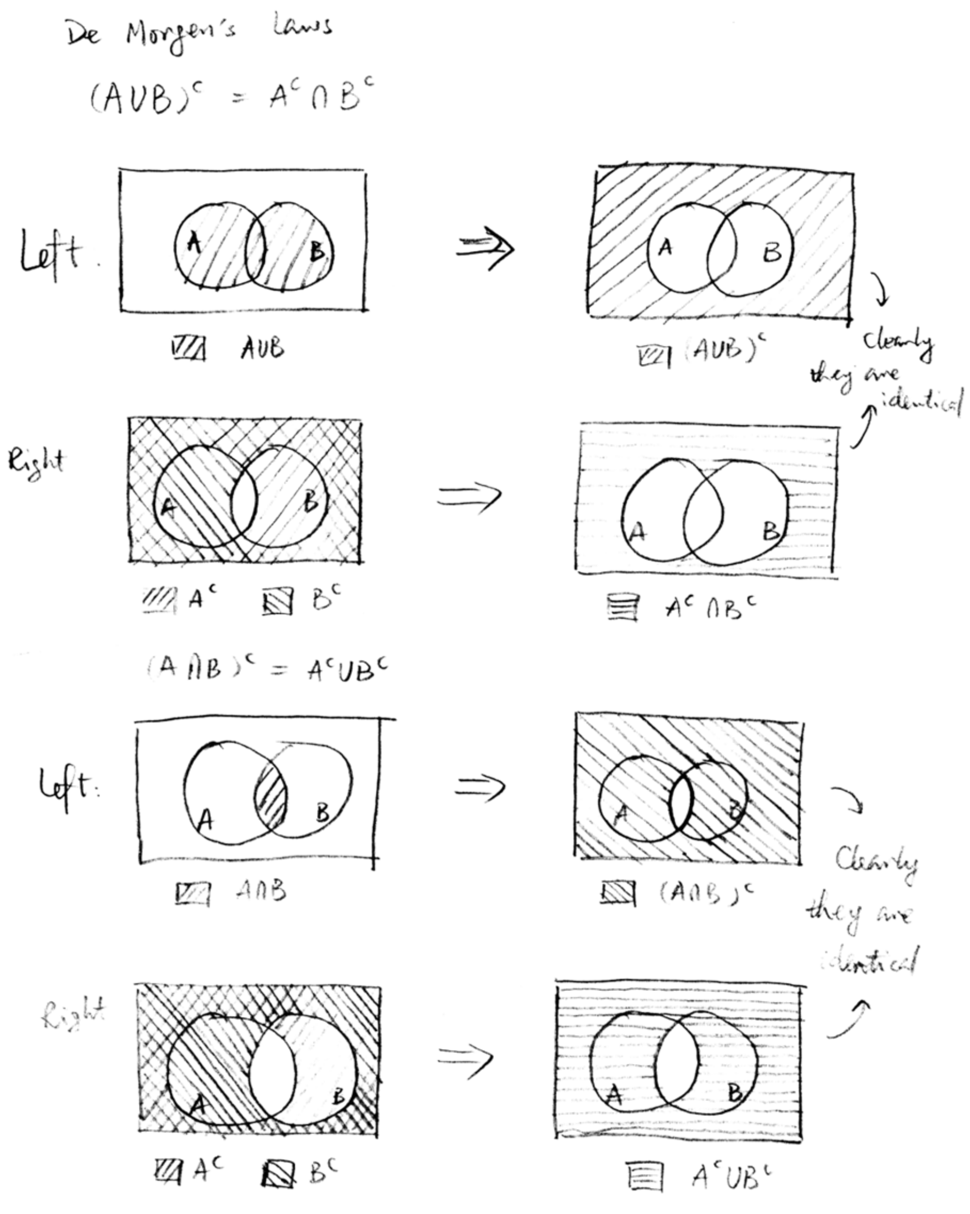
2.3.2 *Identities*

Various equalities can be described through set operations – idempotent law, commutative law, associative law, distributive law, involution law, de Morgan's law etc.

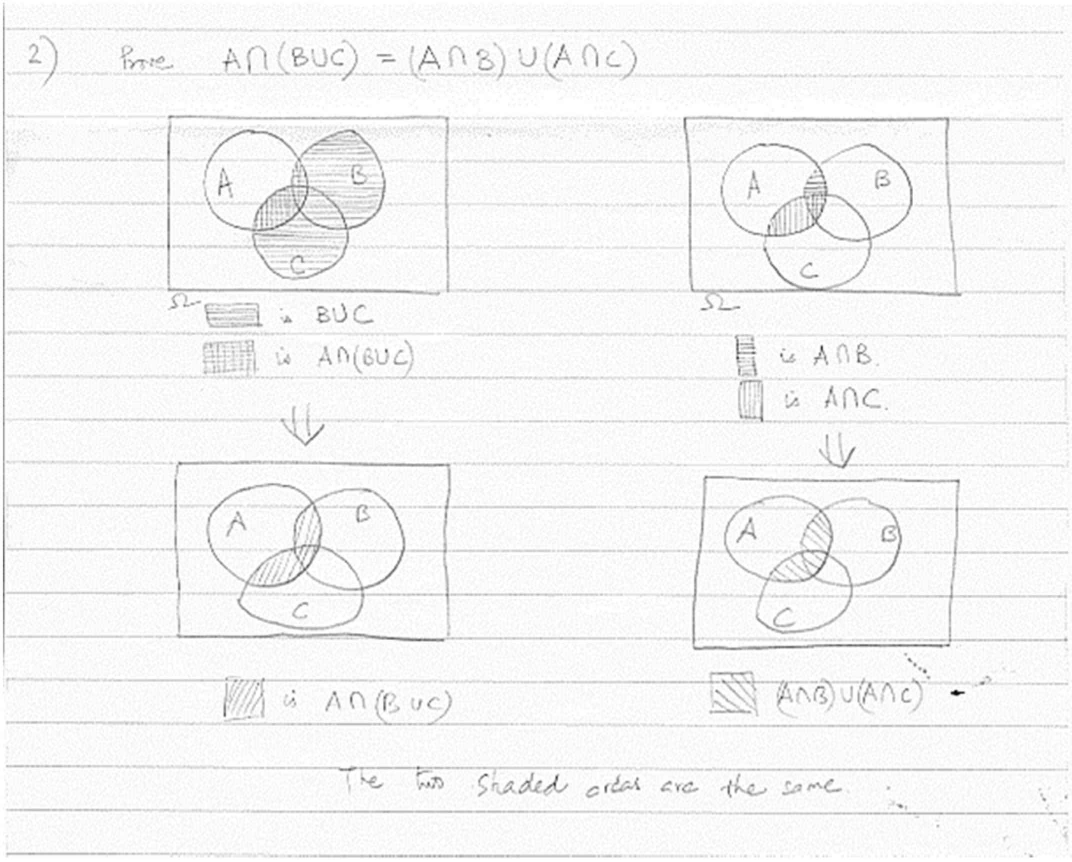
Idempotent laws			
(1*a*)	$A \bigcup A = A$	(1*b*)	$A \bigcap A = A$
Associative laws			
(2*a*)	$(A \bigcup B) \bigcup C = A \bigcup (B \bigcup C)$	(2*b*)	$(A \bigcap B) \bigcap C = A \bigcap (B \bigcap C)$
Commutative laws			
(3*a*)	$A \bigcup B = B \bigcup A$	(3*b*)	$A \bigcap B = B \bigcap A$
Distributive laws			
(4*a*)	$A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$	(4*b*)	$A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$
Identity laws			
(5*a*)	$A \bigcup \varnothing = A$	(5*b*)	$A \bigcap U = A$
(6*a*)	$A \bigcup U = U$	(6*b*)	$A \bigcap \varnothing = \varnothing$
Involution law			
(7)	$(A^{c})^{c} = A$		

	Complement laws		
(8a)	$A \cup A^c = U$	(8b)	$A \cap A^c = \varnothing$
(9a)	$U^c = \varnothing$	(9b)	$\varnothing^c = U$
	DeMorgan`s laws		
(10a)	$(A \cup B)^c = A^c \cap B^c$	(10b)	$(A \cap B)^c = A^c \cup B^c$

2.3.2.1 Prove DeMorgan's Laws



2.3.2.2 Prove the second distributive law



2.3.2.3 Prove (AUB)\(AB) = (A\B)U(B\A)

The left hand side is equal to :

$LHS = (A \cup B) \setminus (AB)$

$= (A \cup B)(AB)^c$ using the identity that $X \setminus Y = XY^c$

$= (A \cup B)(A^c \cup B^c)$ using deMorgan's law

$=[(A \cup B)A^c] \cup [(A \cup B)B^c]$ using distributive law

$=[(AA^c) \cup (BA^c)] \cup [(AB^c) \cup (BB^c)]$ using distributive law

$= \varnothing \cup (BA^c) \cup (AB^c) \cup \varnothing$ where $\varnothing$ is the null set

$= (B \setminus A) \cup (A \setminus B)$ using the identity that $XY^c = X \setminus Y$

$= RHS$

2.3.3 Partition of a set

A partition $\mathfrak{P} = \{A_i\}$ of the universal set U is a collection of mutually exclusive and collectively exhaustive sets, A_i:

$$A_i \cap A_j = \varnothing, \; i \neq j$$
$$\bigcup_i A_i = U \tag{2.3}$$

2.4 Limits of sets

Feller IV.1: Recall limit of the sequence of functions $\{f_i(x)\}$. For a given value of *x* (*x* is omitted to make the notation cleaner), the term lim inf f_n denotes the maximum of the sequence of minima:

$$\liminf_{n\to\infty} f_n = \max_{0<j<\infty}\{\min f_j, f_{j+1}, ...\} = \bigcup_{j=1}^{\infty}\bigcap_{k=j}^{\infty} f_k \tag{2.4}$$

The notation on the right is the shorthand for max (denoted by cup or union) and for min (denoted by cap or intersection), respectively. Likewise, the minimum of the sequence of maxima is:

$$\limsup_{n\to\infty} f_n = \min_{0<j<\infty}\{\max f_j, f_{j+1}, ...\} = \bigcap_{j=1}^{\infty}\bigcup_{k=j}^{\infty} f_k \tag{2.5}$$

These concepts directly apply to a sequence of sets $\{A_k\}$. The infimum and supremum of the sequence are the sets defined respectively as:

$$\liminf_{n\to\infty} A_n = \bigcup_{j=1}^{\infty}\bigcap_{k=j}^{\infty} A_k = \left\{ w : \sum_k (1 - \mathrm{I}_{A_k}(w)) < \infty \right\} \quad \left(\begin{array}{l} \text{i.e., set of points absent in a} \\ \text{finite number of } A_k\text{'s} \end{array} \right) \tag{2.6}$$

$$\limsup_{n\to\infty} A_n = \bigcap_{j=1}^{\infty}\bigcup_{k=j}^{\infty} A_k = \left\{ w : \sum_k I_{A_k}(w) = \infty \right\} \quad \left(\begin{array}{l} \text{i.e., set of points present in} \\ \text{all } A_k\text{'s} \end{array} \right) \tag{2.7}$$

The limit of the sequence $\{A_k\}$ exists if the two limits are equal and may be denoted *A*:

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A \tag{2.8}$$

which may be written in short as:

$$\lim_{n\to\infty} A_n = A \quad \text{or} \quad A_n \to A \tag{2.9}$$

From P6, Resnik Probability Path:

let $A_k = \left[0, \frac{k}{k+1}\right]$

$B_n = \inf_{k \geq n} A_k = \bigcap_{k=n}^{\infty}\left[0, \frac{k}{k+1}\right] = \left[0, \frac{n}{n+1}\right]$

Then, $\liminf_{n\to\infty}\left[0, \frac{n}{n+1}\right] = \bigcup_{n=1}^{\infty}\inf_{k\geq n}\left[0, \frac{k}{k+1}\right] = \bigcup_{n=1}^{\infty}\left[0, \frac{n}{n+1}\right] = [0,1]$

$B^n = \sup_{k\geq n} A_k = \bigcup_{k=n}^{\infty} A_k = \bigcup_{k=n}^{\infty}\left[0, \frac{k}{k+1}\right] = [0,1]$

$\limsup_{n\to\infty} A_n = \lim_{n\to\infty} B^n = \lim_{n\to\infty}[0,1] = [0,1]$

2.5 Ordered sets

An ordered set S is a set in which an order is defined. An order on a set S is a relation "<" with the following two properties (Rudin):

1. If $x \in S$ and $y \in S$ then one and only one of the following three statements is true:

$$x < y \qquad\qquad x = y \qquad\qquad y < x \tag{2.10}$$

2. If three elements *x*,*y* and *z* belong to S, and

$$\text{if } x < y \text{ and } y < z, \text{then } x < z \tag{2.11}$$

The relation "<" may be read as "is less than" or "is smaller than" or "precedes".

2.5.1 *Supremum or least upper bound:*

S is an ordered set and $E \subset S$. If there exists $\beta \in S$ such that $x \leq \beta$ for every $x \in E$, we say E is bounded above and β is an upper bound of E. Now, if α is an upper bound of E such that any $\gamma < \alpha$ is not an upper bound of E, then α is called the least upper bound or supremum of E:

$$\alpha = \sup E \tag{2.12}$$

The supremum α may or may not belong to E.

An ordered set S has the l.u.b. property if for any non-empty subset E that is bounded above, its supremum $\alpha = \sup E$ exists in S.

2.5.2 *Infimum or greatest lower bound:*

Similar to the sup definition above:

Let G be an ordered set and $G \subset S$. If there exists $\beta \in S$ such that $x \geq \beta$ for every $x \in G$, we say G is bounded below and β is a lower bound of G. Now, if α is a lower bound of G such that any $\gamma > \alpha$ is not a lower bound of G, then α is called the greatest lower bound or infimum of G:

$$\alpha = \inf G \qquad (2.13)$$

The infimum α may or may not belong to G.

An ordered set S has the g.l.b. property if for any non-empty subset G that is bounded below, its infimum $\alpha = \inf G$ exists in S.

2.6 Set algebra

Let Ω be any set[1]. A non-empty collection $\mathcal{A}$ of subsets of Ω is an algebra of sets (i.e., a *field*) if: whenever A_1, A_2 are in $\mathcal{A}$, so are $\Omega \backslash A_1$ (i.e., complement of A_1) and $A_1 \cup A_2$ (and therefore A_1A_2 also). Generalizing, if $A_1, A_2, \ldots, A_n$ (n finite) are in $\mathcal{A}$, so are $A_1 \cup A_2 \ldots \cup A_n$ and $A_1A_2 \ldots A_n$.

Example: Let Ω ={a,b,c}. Then we could define a field $\mathcal{A}$ as:

$\mathcal{A} = \phi, \Omega, \{a\}, \{b,c\}$

Example:

S={1,2,3,4}

Describe F, the smallest field containing {1} and {2,3}

$F = \{\{\emptyset\}, \{1\}, \{4\}, \{1,4\}, \{2,3\}, \{2,3,4\}, \{1,2,3,4\}\}$

σ algebra (or σ field or Borel field): The algebra described above is a σ algebra of sets if it holds for a countably infinite[2] collection $A_1, A_2, \ldots$. That is, whenever, the sequence $A_1, A_2, \ldots$, belongs to $\mathcal{A}$, so does $\bigcup_{i=1}^{\infty} A_i$. In other words, a σ algebra $\mathcal{A}$ of subsets of a given set Ω contains the empty set ϕ and is closed with respect to complementation and countable unions.[3]

Measurable space: A couple (Ω, $\mathcal{A}$) is a measurable space where Ω is any set and $\mathcal{A}$ is a σ algebra of subsets of Ω. A subset A of Ω is measurable with respect to $\mathcal{A}$ if $A \in \mathcal{A}$.

Measure: A measure *m* on a measurable space (Ω, $\mathcal{A}$) is a non-negative set function defined for all sets of the σ algebra $\mathcal{A}$, if it has the properties:

---

[1] in the context of probability, X is considered to be the sample space Ω

[2] A **finite sequence** (of size *n*) is a function whose domain is the first *n* natural numbers. An **infinite sequence** is a function whose domain is the set N of natural numbers. A set A is called **countable** if it is the range of some sequence (finite or infinite). The set A is **finite and countable** if it is the range of some finite sequence. The set A is **countably infinite** if it is the range of some infinite sequence.

[3] When the events A1, A2, …, are countably infinite, we can take the Borel field constituting the probability space to consist of *all* subsets of Ω. When Ω is uncountably infinite (e.g., the real line R), we do not want the Borel field to be the collection of *all* subsets of Ω to constitute a probability space. In this case, we only consider events of the type $X \le x_i$ and the Borel field to consist of all finite intervals in R.

(i) $m(\phi) = 0.$ (2.14)

(ii) If $A_1, A_2, \ldots$ is a sequence of disjoint sets of $\mathcal{A}$,

$$\text{then } m(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} m(A_i) . \tag{2.15}$$

Measure space: A measure space (Ω, $\mathcal{A}$, m) means a measurable space (Ω, $\mathcal{A}$) together with a measure *m* defined on $\mathcal{A}$.

Measurable function: Let (Ω, $\mathcal{A}$) and (Ω', $\mathcal{A}$') be two measurable spaces. Then the function (or map) $f : \Omega \to \Omega'$ is called measurable if the inverse satisfies

$$f^{-1}(\mathcal{A}') \subset \mathcal{A} \tag{2.16}$$

In the special case that Ω is the sample space and the range of *f* is the extended real line, i.e., $\Omega' = \mathbb{R}$ and $\mathcal{A}' = B(\mathbb{R})$ the sigma algebra of intervals on the real line, then *f* must satisfy any one of the following conditions in order to be a measurable function with respect to $\mathcal{A}$:

$$\{x : f(x) < \alpha\} \in \mathcal{A} \text{ for each } \alpha$$
$$\{x : f(x) \le \alpha\} \in \mathcal{A} \text{ for each } \alpha$$
$$\{x : f(x) > \alpha\} \in \mathcal{A} \text{ for each } \alpha$$
$$\{x : f(x) \ge \alpha\} \in \mathcal{A} \text{ for each } \alpha$$

Then, as we will see in CHAPTER 4, *f* is called a random variable.