Module 1: Some Basic Concepts

1.1 Review of Set Theory and Notation (*Weiss §1.2*)

► Some Notation and Definitions:

- A **set** is a collection of *distinct* elements (numbers, individuals, objects, ...).
- If A is a set and x is an element of A, we write: $x \in A$.
- If y is **not** an element of A, we write $y \notin A$.
- $\emptyset$ denotes the **empty set** and U the **Universal set**, i.e. the set of all elements of interest.
- The **complement** of A, denoted A^c is

$$A^c = \{x \in U : x \notin A\}.$$

  This is sometimes also denoted A' or $\bar{A}$.
- The set A is a subset of B, denoted $A \subset B$, if all elements of A also belong to B.
  In other words, if $x \in A \implies x \in B$, then $A \subset B$.
- Two sets A and B are equal ($A = B$)

$$\iff A \subset B \text{ and } B \subset A.$$

- The **union** of A and B is

$$A \cup B = \{x \in U : x \in A \text{ or } x \in B\},$$

  i.e. it is the set of all elements belonging to *at least one* of the two sets.
  As a result: $A \cup \emptyset = A$ and $A \cup A^c = U$.
- The **intersection** of A and B is

$$A \cap B = \{x \in U : x \in A \text{ and } x \in B\} = AB,$$

  i.e. it is the set of all elements belonging to *both* sets.
  As a result: $A \cap \emptyset = \emptyset$.
- The sets A and B are **disjoint** or **mutually exclusive** if $A \cap B = AB = \emptyset$,
  i.e. if they have no elements in common.

Proposition 1.1. Commutative Laws.

Let A and B be subsets of U.

1. $A \cup B = B \cup A$
2. $A \cap B = B \cap A$ or $AB = BA$

Proposition 1.2. Associative Laws.

Let A, B and C be subsets of U.

1. $A \cup (B \cup C) = (A \cup B) \cup C = A \cup B \cup C$
2. $A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C$ or $A \cap (BC) = (AB) \cap C = ABC$

Proposition 1.3. Distributive Laws.

Let A, B and C be subsets of U.

1. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
2. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Proposition 1.4. De Morgan's Laws.

Let A and B be subsets of U.

1. $(A \cup B)^c = A^c \cap B^c$ or $(A \cup B)^c = A^c B^c$
2. $(A \cap B)^c = A^c \cup B^c$ or $(AB)^c = A^c \cup B^c$

► Some more notation and definitions:

Let $A_1, A_2, \ldots, A_n$ be subsets of U.

- Then, we write

$$\begin{aligned} A_1 \cup A_2 \cup \ldots \cup A_n &= \bigcup_{i=1}^{n} A_i \\ &= \Big\{ x \in U : x \in A_i \text{ for some } i = 1, 2 \ldots, n \Big\}, \end{aligned}$$

  i.e. the set of x's belonging to *at least one* of $A_1, A_2, \ldots, A_n$.

- We also write

$$\begin{aligned} A_1 \cap A_2 \cap \ldots \cap A_n &= \bigcap_{i=1}^{n} A_i \\ &= \Big\{ x \in U : x \in A_i \text{ for all } i = 1, 2 \ldots, n \Big\}, \end{aligned}$$

  i.e. the set of x's belonging to *all* of $A_1, A_2, \ldots, A_n$.

- The sets $A_1, A_2, \ldots, A_n$ are said to be **pairwise disjoint** or **mutually exclusive** if

$$A_i \cap A_j = \emptyset \quad \text{for } i \neq j.$$

Proposition 1.5. Generalized Distributive Laws.

Let A and $B_1, B_2, \ldots$ be subsets of U.

1. $A \cap \left(\bigcup_{i=1}^{n} B_i \right) = \bigcup_{i=1}^{n} AB_i \qquad \text{for all } n \geq 2$
2. $A \cup \left(\bigcap_{i=1}^{n} B_i \right) = \bigcap_{i=1}^{n} (A \cup B_i) \qquad \text{for all } n \geq 2$

Proposition 1.6. Generalized De Morgan's Laws.

Let $A_1, A_2, \ldots$ be subsets of U.

1. $\left(\bigcup_{i=1}^{n} A_i \right)^c = \bigcap_{i=1}^{n} A_i^c \qquad \text{for all } n \geq 2$
2. $\left(\bigcap_{i=1}^{n} A_i \right)^c = \bigcup_{i=1}^{n} A_i^c \qquad \text{for all } n \geq 2$

► One last definition:

- The sets $A_1, A_2, \ldots, A_n$ form a **partition** of B if
  1. $A_i A_j = \emptyset \quad \text{for all } i \neq j,$ (i.e. they are mutually exclusive),
  2. $\bigcup_{i=1}^{n} A_i = B$.

  Essentially, each A_i is a distinct "piece" of B.

- A and A^c always form a partition of U!

1.2 Sample Space and Events (*Weiss* §*2.1*)

► A few definitions:

- A **random experiment** is an action whose outcome cannot be predicted with certainty.

  Think of:
  - flipping a coin,
  - rolling a die,
  - arrival of patients at the ER in a hospital,
  - playing the lotto 6/49,
  - etc.

- The set of all possible outcomes for a random experiment is called the **sample space**. It is denoted Ω ($\mathcal{S}$ is also used sometimes).

  To denote an individual outcome, we use ω.

- An **event** is a subset of the sample space.

  Events are denoted by capital letters (they are sets).

- A **simple event** is an event consisting of only one outcome.

- An event E is said to occur when the outcome of the random experiment is an element of E.

Example 1.1. Flipping a card from a regular 52 card deck (*cf*. Figure 2.2 from Weiss)

Describe the sample space Ω for this random experiment, and consider the following three events:

D: the card is a diamond,
H: the card is a "high" card (i.e. 10, J, Q, K or A),
F: the card is a "face" card (i.e. J, Q or K).

Also, find at least three partitions of Ω.

Example 1.2. Rolling two dice (*cf*. Figure 2.1 from Weiss)

Describe the sample space of this random experiment.

Are the following events mutually exclusive?

A: the sum is at most 4,
B: at least one result of 4 or more,
C: the difference between the two dice is 3,
D: the sum of the dice is 7.

1.3 Axioms of Probability (*Weiss §2.2, 2.3*)

Definition 1.7. Kolmogorov's Axioms

Let Ω be the sample space of some random experiment.

A function $\mathbb{P}$ defined on the events of Ω is called a **probability measure** if it satisfies the following three properties.

1. The **nonnegativity axiom**:
$$\mathbb{P}(E) \geq 0 \qquad \text{for all events } E \subset \Omega.$$
2. The **certainty axiom**:
$$\mathbb{P}(\Omega) = 1,$$
i.e. one of the possible outcomes will occur.
3. The **additivity axiom**:
If $A_1, A_2, \ldots$ are mutually exclusive subsets of Ω, then
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Proposition 1.8. Two consequences of Kolmogorov's axioms

1. For the empty set, we have $\mathbb{P}(\emptyset) = 0$.
2. Finite additivity:
Let $A_1, A_2, \ldots$ be mutually exclusive events of Ω. Then
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) \qquad \text{for all } n \in \mathbb{N}.$$

Proposition 1.9. Finite or infinite countable sample space

For any event $E \subset \Omega$,

$$\mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\{\omega\}),$$

i.e. $\mathbb{P}(E)$ is equal to the sum of the probabilities of each outcome in E. Also,

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

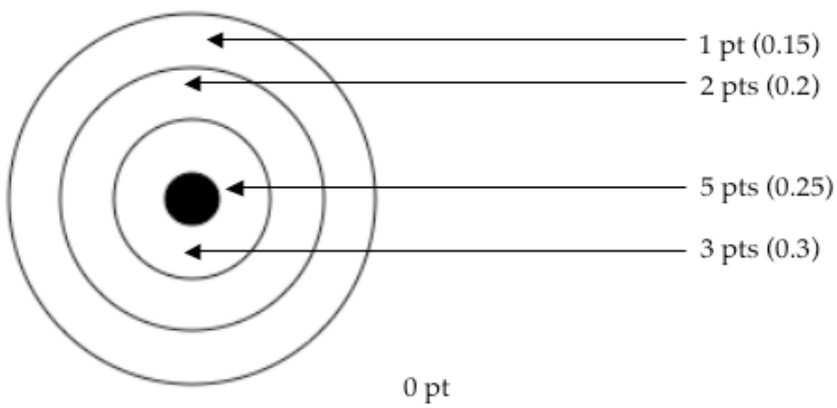
i.e. the probabilities of all the outcomes in Ω sum to one.

► Conclusion:

We can assign probabilities to each outcome and calculate the probabilities of events by adding the probabilities of the individual outcomes!

Example 1.3. Archery

An archer fires at a target.



Define the following events:

A: the arrow lands inside the inner circle (i.e. the shot scores 3 pts or more),
M: the arrow misses the target completely.

Find $\mathbb{P}(A)$ and $\mathbb{P}(M)$.

Proposition 1.10. Equal-likelihood model or classical probability model

Let Ω be a finite sample space with equally likely outcomes. Then, for any event $E \subset \Omega$, we have

$$\mathbb{P}(E) = \frac{N(E)}{N(\Omega)} = \frac{\text{nb. of outcomes in } E}{\text{nb. of outcomes in } \Omega}.$$

► Note:

The previous result implies that in the case of a finite sample space,

$$0 \leq \mathbb{P}(E) \leq 1 \quad \text{for any } E \subset \Omega, \qquad \text{and} \qquad \mathbb{P}(\emptyset) = 0.$$

Example 1.4. (*cf*. Example 1.2 on page 4)

When rolling two dice,

$$\Omega = \Big\{(x, y) : x, y \in \{1, 2, \ldots, 6\}\Big\}.$$

Assume that the $N(\Omega) = 36$ outcomes are equally likely. Find $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(C)$ and $\mathbb{P}(D)$.

Example 1.5. Flipping a coin 3 times

What is the probability of seeing a sequence of at least 2 consecutive identical results?

1.4 General Properties of Probability (*Weiss* §*2.4*)

Proposition 1.11. Complementation Rule

Let $A \subset \Omega$. Then,

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

Proposition 1.12. Domination Principle

If $A \subset B \subset \Omega$ are any events, then

$$\mathbb{P}(A) \leq \mathbb{P}(B).$$

► Two consequences of the domination principle:

- For any $E \subset \Omega$, we have that $\mathbb{P}(E) \leq 1$.
- For any $A, B \subset \Omega$,

$$\mathbb{P}(AB) \leq \mathbb{P}(A) \leq \mathbb{P}(A \cup B) \qquad \text{and} \qquad \mathbb{P}(AB) \leq \mathbb{P}(B) \leq \mathbb{P}(A \cup B).$$

Example 1.6. (*cf*. Example 1.4)

Argue that $D \subset B$ and verify that $\mathbb{P}(D) \leq \mathbb{P}(B)$. Also, find the probability that the sum of the dice is not 7.

Proposition 1.13. General Addition Rule for two events
Let $A, B \subset \Omega$. Then,
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB).$$

► Proving the above property, we have obtained the following result that is often useful:
$$\mathbb{P}(A^c B) = \mathbb{P}(B) - \mathbb{P}(AB).$$

Example 1.7. Defective CD's
Of all CD's produced by a manufacturer,

- 3% have a surface defect,
- 8% have a balance defect,
- 91% are defect-free.

Find the probability a randomly selected CD has:

- at least one type of defect,
- both defects,
- only a surface defect,
- only a balance defect.

Proposition 1.14. Inclusion-exclusion principle
Let $A_1, A_2, \ldots$ be events (and subsets of Ω). Then, for all $n \geq 2$,
$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i_1<i_2}\sum \mathbb{P}(A_{i_1}A_{i_2}) + \sum_{i_1<i_2<i_3}\sum\sum \mathbb{P}(A_{i_1}A_{i_2}A_{i_3}) + \cdots$$
$$+ (-1)^{k+1} \sum_{i_1<i_2<\cdots<i_k}\sum\sum \mathbb{P}(A_{i_1}A_{i_2}\cdots A_{i_k}) + \cdots + (-1)^{n+1}\mathbb{P}(A_1A_2\cdots A_n).$$

In other words, the probability of the union is equal to

- the sum of the probabilities of each event,
- $-$ the sum of the probabilities of all 2×2 (pairwise) intersections,
- $+$ the sum of the probabilities of the 3×3 intersections,
- etc.

Example 1.8. (*cf*. Example 1.4 on page 7)

We have defined:

A: the sum is at most 4,
D: the sum of the dice is 7,
E: there is at least one result of 3.

Calculate $\mathbb{P}(A \cup D \cup E)$.

Example 1.9. Wrong numbers?

We are told that, in a study of 1000 subscribers to a certain magazine, the following data was reported:

P:	Professionals,	$N(P) = 312$,	$N(MP) = 86$,	$N(MPC) = 25$.
M:	Married,	$N(M) = 470$,	$N(PC) = 42$,	
C:	College graduates,	$N(C) = 525$,	$N(MC) = 147$,	

How can we tell that there is an error in the reported data?