Experiment: Any procedure that can be repeated (at least hypothetically) an infinite times and has a well-defined set of possible outcomes.

Sample Space: The collection of all possible outcomes of an experiment is called the sample space of the experiment. In this class we use S to denote sample space.

Event: An event is a subset of the sample space S.

Examples of experiment and sample space

Ex1: If the experiment is flipping a coin, sample space is $S = \{H, T\}$.

Ex2: If we flip two coins, sample space is $S = \{(H, H), (H, T), (T, H), (T, T)\}$.

Ex3: In Ex2, if we are only interested in the number of Heads, sample space is $S = \{2, 1, 0\}$.

Ex4: Pick a die from a black box containing three dices: red, green and yellow, $S = \{R, G, Y\}$.

Ex5: Roll a die, $S = \{1, 2, 3, 4, 5, 6\}$.

Ex6: Toss a coin until see a Head. $S = \{H, TH, TTH, TTTH, TTTTH, T \cdots TH, \ldots\}$. Note that S has infinite many outcomes.

Some examples of Events

Ex7: In the example of flipping two coins $S = \{(H,H),(H,T),(T,H),(T,T)\}$.

Consider $A = \{(H,H),(H,T)\}$. Then A is an event. Note: A is the event that the first coin shows "Head".

Consider $A_2 = \{(T,H)\}$. A_2 is another event.

Consider $A_3 = \{(H,H),(H,T),(T,T)\}$. A_3 is the event containing the outcomes that are not in event A_2.

Set Algebra

Relations:

1. **Inclusion:** an event A is contained in another event B, if every outcome in A also belongs to B. We write $A \subset B$, if A is contained in B.

2. **Equality:** we say $A = B$ if $A \subset B$ and $B \subset A$. Note in order to prove two events (sets) A and B are the same, we often check $A \subset B$ and $B \subset A$.

3. **Complementation:** the complement of an event A is defined to be the collection of all outcomes in sample space S that do not belong to A. We write it A^c. E.g., In Ex7 $A_2^c = A_3$ and $A_3^c = A_2$.

Operations

1. **Intersection:** the intersection of two events A and B, denoted by $A \cap B$, is the collection of all outcomes that belong to BOTH A AND B.

2. **Union:** the union of two events A and B, denoted by $A \cup B$, is the collection of all outcomes that belong to EITHER A OR B (OR BOTH).

3. **Empty set:** If there is no common outcome in A and B, then we say $A \cap B = \emptyset$.

Examples

Ex1: $A \cap A^c = \emptyset$; $A \cap S = A$; $A \cup A^c = S$

Ex2: Roll a die. $S = \{1, 2, 3, 4, 5, 6\}$. Let $A = \{1, 3, 5\}$ and $B = \{1, 2, 3\}$.

Then $A^c = \{2, 4, 6\}$ and $B^c = \{4, 5, 6\}$.

$A \cap B = \{1, 3\}$ and $A \cup B = \{1, 2, 3, 5\}$.

$(A \cap B)^c = \{2, 4, 5, 6\}$ and $(A \cup B)^c = \{4, 6\}$.

$A^c \cap B^c = \{4, 6\}$ and $A^c \cup B^c = \{2, 4, 5, 6\}$.

We see $(A \cap B)^c = A^c \cup B^c$ and $(A \cup B)^c = A^c \cap B^c$. This is always true! (Can you prove it?)

More than 2 events

Suppose we have a sequence of events, denoted by $A_1, A_2, \ldots, A_n$. For example, $n = 60$.

Intersection of these n events is denoted by $\cap_{i=1}^{n} A_i$ which contains all outcomes that are common to all these n events.

Union of these n events is denoted by $\cup_{i=1}^{n} A_i$ which contains all outcomes that belong to at least one of the n events.

Disjoint Events: We say a sequence of events $\{A_i\}$ $i = 1, 2, \ldots$ are disjoint events, if $A_i \cap A_j = \emptyset$ for each pair of (i, j), $i \neq j$.

Three Axioms of Probability:

Suppose we have a system to assign a number to each event in the sample space S. Let $P(A)$ be the number assigned to event A. A Probability model on a sample space S is a specification of numbers to the events in S such that the following three Axioms are satisfied

- Axiom 1. For every event A, $P(A) \geq 0$.
- Axiom 2. $P(S) = 1$.
- Axiom 3. If $\{A_i\}$ $i = 1, 2, \ldots$ are disjoint events then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Finite Sample Spaces

Definition: Consider experiments for which the sample space S contains only a finite number of points $s_1, \dots, s_n$. A probability distribution on S is specified by assigning a probability p_i to each point s_i for each $i = 1, 2, \dots, n$. p_i is called the probability that s_i will be the outcome of the experiment. The numbers $p_1, p_2, \dots, p_n$ must satisfy two conditions:

(1) $p_i \geq 0$ for all i

(2) $\sum_{i=1}^{n} p_i = 1$.

The probability of each event A can be found by adding the probability p_i of all s_i that belongs to A.

For equally likely outcomes, we let $p_i = \frac{1}{n}$ for all $i = 1, 2, \dots, n$.

Examples

Ex1: Flip a coin. $S = \{H, T\}$. $s_1 = H$ and $s_2 = T$. Let $p_1 = p_2 = 0.5$.

Or consider a biased coin. $p_1 = 0.55$ and $p_2 = 0.45$.

Ex2: Roll a die. $S = \{1, 2, 3, 4, 5, 6\}$ We can let $p_i = \frac{1}{6}$ for $i = 1, 2, 3, 4, 5, 6$. This is the probability model for throwing a die.

Ex3: Roll two dice. $S = \{(x, y), 1 \leq x \leq 6; 1 \leq y \leq 6\}$. x is the number on the first die and y is the number on the second die. Let A be the event that $x + y = 5$. What is $P(A)$?

Examples, contd'

Solution:

$$
\begin{aligned}
S = \; & (1,1),(1,2),(1,3),(1,4),(1,5),(1,6) \\
& (2,1),(2,2),(2,3),(2,4),(2,5),(2,6) \\
& (3,1),(3,2),(3,3),(3,4),(3,5),(3,6) \\
& (4,1),(4,2),(4,3),(4,4),(4,5),(4,6) \\
& (5,1),(5,2),(5,3),(5,4),(5,5),(5,6) \\
& (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)
\end{aligned}
$$

So we find $A = \{(1,4),(4,1),(2,3),(3,2)\}$ so $P(A) = 1/36 + 1/36 + 1/36 + 1/36 = 1/9$.

Ex4: A card is dealt from a well-shuffled deck of cards. What is the chance of getting a Jack?

Solution: A well-shuffled deck of cards means that each one of the 52 cards has a probability 1/52 to be selected. In another word we assume this is an equally likely outcome model. The event is $A = \{select\ a\ Jack\}$. There are 4 Jacks and each has probability 1/52, so the probability of getting a Jack is $P(A) = 4/52$.

A List of Theorems

Theorem 1.5.1: $P(\emptyset) = 0$. The probability of the empty set is always zero.

Proof: Let $A_i = \emptyset$ for $i = 1, 2, \ldots$. We have a sequence of disjoint events because for any pair of (i, j) $A_i \cap A_j = \emptyset \cap \emptyset = \emptyset$. So we use Axiom 3 to get

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} P(\emptyset)$$

but $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} \emptyset = \emptyset$, so the above equation becomes

$$P(\emptyset) = \sum_{i=1}^{\infty} P(\emptyset)$$

which means $P(\emptyset) = 0$.

Theorem 1.5.2: For every n disjoint events, $A_1, A_2, \ldots, A_n$,

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i).$$

Proof: Let's add $A_{n+i} = \emptyset$ for $i = 1, 2, \ldots$ and consider the infinite sequence $(A_1, \ldots, A_n, \emptyset, \emptyset, \ldots)$. Note that the augmented sequence are disjoint events. Therefore we can use Axiom 3 to write down

$$\begin{aligned} & P(A_1 \cup \cdots \cup A_n \cup \emptyset \cup \emptyset \cup \cdots) \\ = \; & P(A_1) + \cdots + P(A_n) + P(\emptyset) + P(\emptyset) + \cdots \\ = \; & P(A_1) + \cdots + P(A_n) + 0 + 0 + \cdots \\ = \; & P(A_1) + \cdots + P(A_n). \end{aligned}$$

But $A_1 \cup \cdots \cup A_n \cup \emptyset \cup \emptyset \cup \cdots = A_1 \cup A_2 \cdots \cup A_n$. So $P(A_1 \cup A_2 \cdots \cup A_n) = \sum_{i=1}^n P(A_i)$.

Special case: Consider A and B. If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

Theorem 1.5.3: For every event A, $P(A^c) = 1 - P(A)$.

Proof: Note that $A \cap A^c = \emptyset$, so we can use Theorem 1.5.2 (special case) and get

$$P(A \cup A^c) = P(A) + P(A^c)$$

but $A \cup A^c = S$ so $P(A \cup A^c) = P(S) = 1$ (Axiom 2). Thus

$$1 = P(A) + P(A^c)$$

or $P(A^c) = 1 - P(A)$.

This is a very handy formula. In many cases A can be a very complicated event, so it is not easy to compute its probability directly. Quite often, A^c is a much simpler event to think of and $P(A^c)$ is easier to get. So we use $P(A) = 1 - P(A^c)$ to compute $P(A)$.

Theorem 1.5.4: If $A \subset B$ then $P(A) \leq P(B)$.

Proof: Note if $A \subset B$ then $B = A \cup (B \cap A^c)$, A and $(B \cap A^c)$ are disjoint. So $P(B) = P(A) + P(B \cap A^c)$ which means $P(B) - P(A) = P(B \cap A^c)$ and $P(B \cap A^c) \geq 0$ (Axiom 1).

Theorem 1.5.5: For each event A, $0 \leq P(A) \leq 1$.

Proof: By Axiom 1, $0 \leq P(A)$. Note $A \subset S$ and $P(S) = 1$, so by Theorem 1.5.4, $P(A) \leq 1$.

Note: Theorems 1.5.1, 1.5.2, 1.5.4, 1.5.5 are intuitively correct. You only need to know these facts.

Theorem 1.5.6: For every two events A and B, we have

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

Proof: Note

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

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

The three events $(A \cap B^c), (A \cap B), (B \cap A^c)$ are disjoint events, therefore we know the probability of their union must equal to the sum of their probabilities. But their union is exactly $A \cup B$. So we can write

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(B \cap A^c) \quad (1).$$

On the other hand, we apply Theorem 1.5.2 to both A and B,

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

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

$$P(A) + P(B) = P(A \cap B^c) + 2P(A \cap B) + P(B \cap A^c)$$

$$\begin{aligned} P(A) + P(B) - P(AB) &= P(A \cap B^c) + 2P(A \cap B) \\ &\quad + P(B \cap A^c) - P(A \cap B) \end{aligned}$$

(2).

You can now compare (1) and (2) and prove the theorem.

Note: Theorem 1.5.6 covers the case when A and B are disjoint. Then $P(A \cap B) = P(\emptyset) = 0$, the equation in Theorem 1.5.6 reduces to

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

Ex1. Consider two events: A and B. Let $P(A) = 1/3$ and $P(B) = 1/2$. Determine the value of $P(BA^c)$ for each of the following conditions (a). A and B are disjoint; (b) $A \subset B$; (c) $P(AB) = 1/8$.

Solution: (a) If A and B are disjoint, then B is contained in A^c. So $BA^c = B \cap A^c = B$ which means $P(BA^c) = P(B) = 1/2$.

(b) If $A \subset B$ then note $B = (BA) \cup (BA^c) = A \cup BA^c$ and $A \cap BA^c = \emptyset$. So $P(B) = P(A) + P(BA^c)$ which means $P(BA^c) = P(B) - P(A) = 1/2 - 1/3 = 1/6$.

(c) note $B = (BA) \cup (BA^c)$ and $(BA) \cap (BA^c) = \emptyset$. So $P(B) = P(BA) + P(BA^c)$ which means $P(BA^c) = P(B) - P(BA) = 1/2 - 1/8 = 3/8$.

Ex2. A patient with a sore throat and low grade fever is believed to have either bacterial infection or viral infection. The probability of bacterial infection is 0.7 and the probability of viral infection is 0.4. What is the probability that the patient has both bacterial and viral infection?

Solution: Let BI denote the event of bacterial infection and VI denote the event of viral infection. Then $P(BI) = 0.7$ and $P(VI) = 0.4$. The statement "the patient is believed to have either bacterial infection or viral infection" means that $P(BI \cup VI) = 1$. On the other hand, $P(BI \cup VI) = P(BI) + P(VI) - P(BI \cap VI)$, so $P(BI \cap VI) = 0.7 + 0.4 - 1 = 0.1$. But $BI \cap VI$ is the event that the patient has both bacterial and viral infection. So the desired probability is 0.1.

Ex3. A student applies to graduate school. He only applies two universities X and Y. He estimates that his probability of being accepted by X is 0.7, and by Y is 0.4. He also suspects that there is a 75 chance that at least one school will reject him. Then what is the probability that at least one of the universities will accept him?

Solution: Let AX denote the event that he is accepted by X, and AY be the event that he is accepted by Y. We know $P(AX) = 0.7$ and $P(AY) = 0.4$. The event that one of the universities will accept him can be written as $AX \cup AY$. So we need to compute $P(AX \cup AY)$. We use $P(AX \cup AY) = P(AX) + P(AY) - P(AX \cap AY)$. If we know $P(AX \cap AY)$, we can compute $P(AX \cup AY)$. For that, note $AX \cap AY$ means that he is accepted by both schools, so $(AX \cap AY)^c$ is the event that at least one school rejects him. We know its probability is 0.75. Thus $P(AX \cap AY) = 1 - P((AX \cap AY)^c) = 1 - 0.75 = 0.25$. So $P(AX \cup AY) = 0.7 + 0.4 - 0.25 = 0.85$.